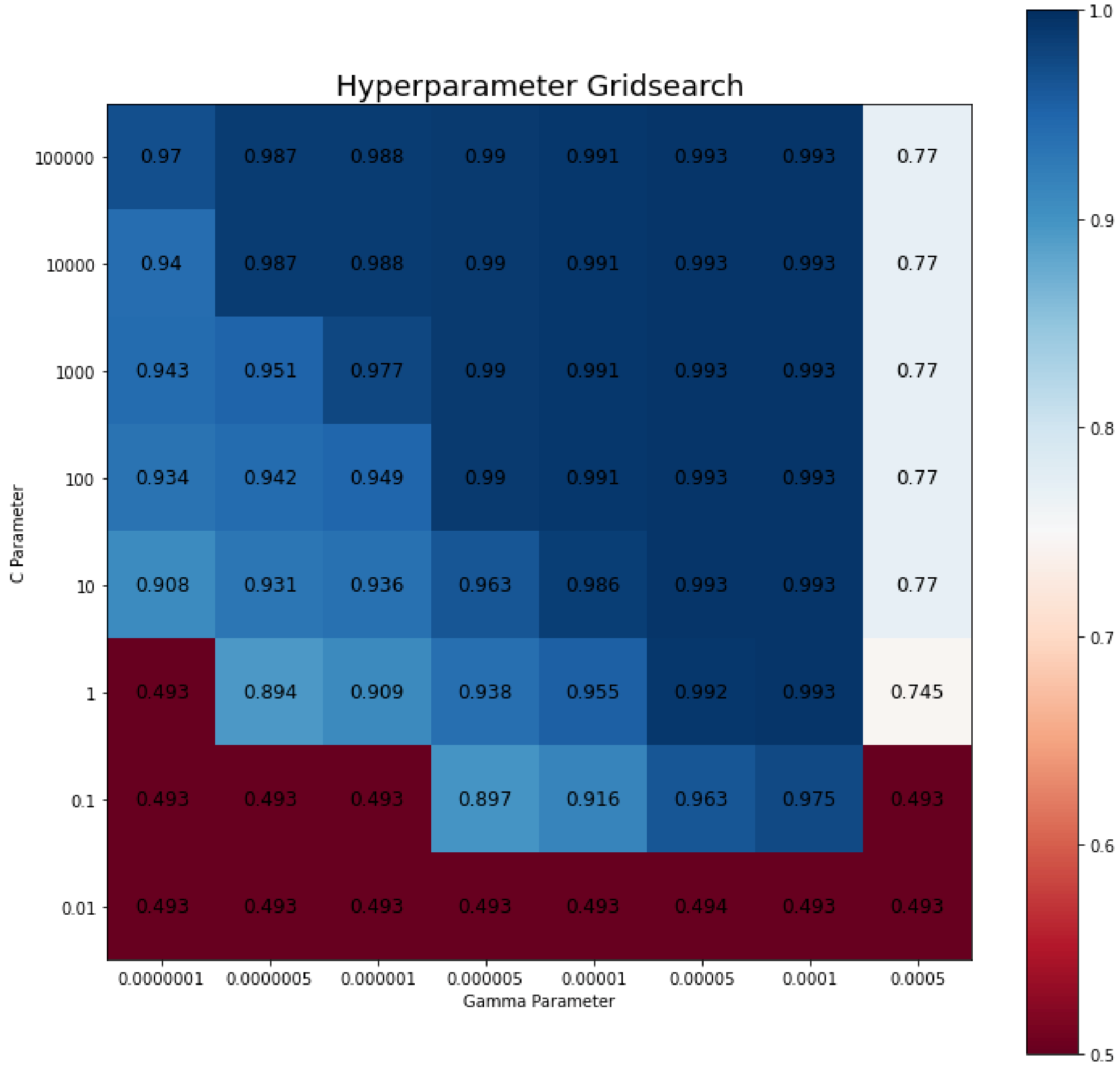
Introduction to Machine Learning HW4

109550159 李驊恩

Coding Part

Result of Grid Search:




```
Best score:
0.9934285714285714
Best gamma:
0.0001
Best C:
1
```

Question Part

1.

1.

$\phi(x_i)^T\phi(x_j) = \phi(x_j)^T\phi(x_i)$ in K . for $i,j = 1 \ldots N$

$\rightarrow$ K is symmetric

Then we do SVD decomposition $\rightarrow K = V\Lambda V^T$ where V is an orthonormal matrix V_t and the diagonal matrix Λ contains the eigenvalues λ_t of K

If K is positive semidefinite, all λ are non-negative.

We can consider the feature map and generate mapping function:

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

The input x_i is projected into an n-dimension space, and the t-th dimension of the i-th piece of data is the i-th dimension of the t-th eigenvector. Only when all of the eigenvalues are positive that we can find the square root of λ_t and thus we can project x_i into n-dimension space.

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

which is the original kernel function

Thus, K should be positive semidefinite is the necessary and sufficient condition for $K(x,x')$ to be a valid kernel.

2.

2.

The Taylor series of $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$

$\Rightarrow k(x,x') = \exp(k_1(x,x')) = 1 + \frac{k_1(x,x')}{1!} + \frac{[k_1(x,x')]^2}{2!} + \cdots = \sum_{n=0}^{\infty} \frac{[k_1(x,x')]^n}{n!}$

Let $q(x) = e^x$, the Taylor series of e^x can be considered as a polynomial whose coefficients is $\frac{1}{n!}$ at the n-power term.

Since $\frac{1}{n!}$ is always positive for $n = 0, 1, 2 \cdots$, $q(x)$ is a non-negative-coefficient polynomial function. Thus, we can apply lemma 6.15 in the slide: Given a valid kernel function $k_1(x,x')$, $k(x,x') = q(k_1(x,x'))$ is also a valid kernel function when $q(\cdot)$ is a polynomial function with non-negative coefficient.

$\Rightarrow k(x,x') = \exp(k_1(x,x')) = q(k_1(x,x'))$

$\Rightarrow k(x,x')$ is also a valid kernel.

※

3.

3.

(a) Let $q(x) = x+1$

$\Rightarrow q(K_1(x,x')) = K_1(x,x')+1$

Given a valid kernel function $K_1(x,x')$, $K(x,x') = q(K_1(x,x'))$ is also a valid kernel function when $q(\cdot)$ is a polynomial function with nonnegative coeffizient. (By 6.15)

Since $q(x) = x+1$ is a non negative-coffizient-polynomial function,

$K_1(x,x')+1 = q(K_1(x,x')) = K(x,x')$

Thus, $K(x,x')$ is valid kernel.

(b) Let $k_1(x,x') = (x^T x')^2$ (slide p.13: $k(x,z) = (x^T z)^2$ is a valid kernel)

$(x_1, x_2) = (1, 0)$, $(x_1', x_2') = (0, 1)$

$\rightarrow k(x,x') = (x^T x')^2 - 1 \rightarrow K = \begin{bmatrix} (1\times1+0\times0)^2-1 & (1\times0+0\times1)^2-1 \\ (1\times0+0\times1)^2-1 & (0\times0+1\times1)^2-1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

Compute eigenvalue of K: $\det(K-\lambda I) = \begin{vmatrix} -\lambda & -1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda+1)(\lambda-1) = 0$

$\Rightarrow \lambda = 1, -1 \Rightarrow$ there is negative eigenvalue $\Rightarrow$ K is not positive semidefinite

$\rightarrow K(x,x')$ is not a valid kernel

#

3.

(C)

We can consider a mapping function: $\phi: x_i \mapsto (e^{x_i^T x_i}, 0)$

Let $k_2(x, x')$ be a valid kernel: $k_2(x,x') = \phi(x)^T \phi(x')$

$\rightarrow \phi(x)^T \phi(x') = (e^{x^T x}, 0) \cdot (e^{x'^T x'}, 0) = e^{\|x\|^2} \cdot e^{\|x'\|^2}$

where $\phi(x)^T = e^{\|x\|^2}$, $\phi(x') = e^{\|x'\|^2}$

$\rightarrow k(x, x') = k_1(x,x')^2 + k_2(x,x')$

By lemma 6.18 in p.15: $k(x,x') = k_1(x,x')k_2(x,x')$ is also valid if $k_1(x,x')$ and $k_2(x,x')$ is valid

$\rightarrow k_1(x,x')^2$ is valid

By lemma 6.17 in p.15: $k(x,x') = k_1(x,x') + k_2(x,x')$ is also valid if $k_1(x,x')$ and $k_2(x,x')$ is valid

$\rightarrow k_1(x,x')^2 + k_2(x,x')$ is valid

$k(x,x') = k_1(x,x')^2 + k_2(x,x') \Rightarrow k(x,x')$ is valid

#

3.

(d)

$$e^{x} = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \text{(Taylor Series)}$$

$$\Rightarrow \exp(K_1(x,x')) = 1 + K_1(x,x') + \frac{1}{2!}K_1(x,x')^2 + \cdots = \sum_{n=0}^{\infty} \frac{(K_1(x,x'))^n}{n!}$$

$$\Rightarrow k(x,x') = K_1(x,x')^2 + \exp(K_1(x,x')) - 1$$

$$= K_1(x,x')^2 + (\cancel{1} + K_1(x,x') + \frac{1}{2!}K_1(x,x')^2 + \cdots) - \cancel{1}$$

$$= K_1(x,x') + \frac{3}{2}K_1(x,x')^2 + \sum_{n=3}^{\infty} \frac{[K_1(x,x')]^n}{n!}$$

Let $q(x) = x + \frac{3}{2}x^2 + \sum_{n=3}^{\infty} \frac{[K_1(x,x')]^n}{n!}$

Given a valid kernel function $K_1(x,x')$, $K(x,x') = q(K_1(x,x')$ is also a valid kernel function when $q(\cdot)$ is a polynomial function with non-negative coefficient. (By 6.15)

Since $q(x) = x + \frac{3}{2}x^2 + \sum_{n=3}^{\infty} \frac{[K_1(x,x')]^n}{n!}$ is a non negative-coefficient polynomial function.

$k(x,x') = K_1(x,x')^2 + \exp(K_1(x,x')) - 1 = K_1(x,x') + \frac{3}{2}K_1(x,x')^2 + \sum_{n=3}^{\infty} \frac{[K_1(x,x')]^n}{n!}$

$= q(K_1(x,x'))$ Thus, $k(x,x')$ is valid kernel.

4.

4.

minimize $(x-2)^2$

subject $(x+3)(x-1) \le 3 \rightarrow x^2+2x-6 \le 0$

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

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

$= (1+\lambda)x^2 + (2\lambda-4)x + (4-6\lambda) \quad (\lambda \ge 0)$

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

Then it become a maximum margin problem:

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

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

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

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

$\rightarrow$ The dual problem: maximize $\frac{-7\lambda^2+2\lambda}{1+\lambda}$

subject to $\lambda \ge 0$

#