Surabhi
2016271

Machine Learning
Assignment 2.

Date

Q2 (Bonus)

The rbf kernel or the Gaussian kernel has the form:

$$K(x, x') = \exp\left(\frac{-\|x - x'\|^2}{2\sigma^2}\right)$$

or more generally

$$\exp\left(-\gamma \|x - x'\|^2\right)$$

gamma

scaling factor $\propto 1/\text{numfeatures}$

The rbf kernel has the capability to map the data to a higher dimensional space by mapping each data point to its own dimension. As number of training samples increase, the dimension also increase.

Overfitting however, is determined not by the size of dataset or the dimension but instead by the parameter γ. γ determines how complex the decision boundary will be. ∴ A higher γ can lead to overfitting while a lower γ can even lead to underfitting. The optimum γ is nearly $1/\text{numfeatures}$.

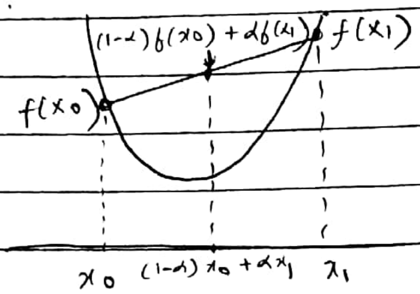
∴ Overfitting can be prevented by choosing a suitable γ value and hence, simply the kernel or dataset size cannot cause overfitting.

Date

Q3. A convex function $f : D \to R$ satifies the inequality:

- $f((1-\alpha)x_0 + \alpha x_1) \leq (1-\alpha) f(x_0) + \alpha f(x_1)$

Visually:



Assume g and h are 2 convex functions.

$\therefore$ we have:

$g((1-\alpha_1)x_0 + \alpha_1 x_1) \leq (1-\alpha_1) g(x_0) + \alpha_1 g(x_1)$

and

$h((1-\alpha_1)x_0' + \alpha_1 x_1') \leq (1-\alpha_1) h(x_0') + \alpha_1 h(x_1')$

Let $j = g + h$. To show: j is a convex fn:

$$
\begin{aligned}
j((1-\alpha_1)x_0 + \alpha_1 x_1) &= g((1-\alpha_1)x_0 + \alpha_1 x_1) \\
&\quad + h((1-\alpha_1)x_0 + \alpha_1 x_1) \\
&\leq (1-\alpha_1) g(x_0) + \alpha_1 g(x_1) \\
&\quad + (1-\alpha_1) h(x_0) + \alpha_1 h(x_1) \\
&= (1-\alpha_1)(g(x_0) + h(x_0)) \\
&\quad + \alpha_1 (g(x_1) + h(x_1))
\end{aligned}
$$

$$
\therefore g((1-\alpha_1)x_0 + \alpha_1 x_1) + h((1-\alpha_1)x_0 + \alpha_1 x_1) \leq (1-\alpha_1)(g(x_0) + h(x_0)) + \alpha_1 (g(x_1) + h(x_1))
$$

$\therefore$ sum of 2 convex fns is also convex

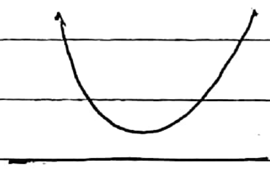
Date

Loss function for L1 Regularized Linear Regression:

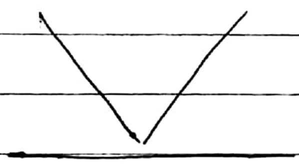
$$L = \underbrace{\sum_{\text{over all samples}} (y_i - f(x_i, w))^2}_{1} + \underbrace{\sum_{w_j \in W} |w_j|}_{2}$$

L is convex if 1 and 2 are convex since sum of convex is convex.

1 is a summation of squares.
since ~~pol~~ quadratic equations are convex:



and absolute function or mod function is convex:



∴ sum of mod functions, and sum of quadratic functions and hence sum of these is also convex.

∴ L is convex