Exercise 5

Question 1 (5pt)

In support vector regression model, what is the relation, if any, between ϵ and noise variance? What is the effect of using different ϵ on overfitting and underfitting?

Question 2 (7pt)

Derive the dual kernel form of support vector regression from its primal form, given as:

$$\begin{aligned} &\min_{w,b} \frac{1}{2}||w||^2 + C\sum_{i=1}^{m}(\hat{\xi}^{(i)} + \xi^{(i)}) \\ &s.t. \quad y^{(i)} - (w\Phi(x^i) + b) \leq \epsilon + \hat{\xi}^{(i)} \\ &\qquad (w\Phi(x^i) + b) - y^{(i)} \leq \epsilon + \xi^{(i)} \\ &\qquad \hat{\xi}^{(i)} \geq 0, \xi^{(i)} \geq 0, \forall i = 1, \ldots, m. \end{aligned} \tag{1}$$

Hint: Chapter 13 in text book "Introduction to Machine Learning".