Q.1.2) c) ii). [Plots were on the previous page].

→ Variation of sigma.

↳ Sigma basically represents the neighborhood that we take around our target point. So, a larger sigma implies a bigger neighbourhood.

→ For sigma = 1, the neighborhod is really small, and as we can see, the model is overfitting on the data (High Variance)

→ For sigma = 100, the neighborhood is really large and the model is under-fitting on the data (High bias).

↳ For sigma = 10, We observe that the model fits just perfectly to the data.

→ Variation of lambda.

→ Higher the lambda, more will be the regularisation penalty, and the variance of the model will decrease (more bias).

↳ At lambda = 0.1, the model fits the data just perfectly, but as we increase lambda, the regularisation penalty forces the model to decrease it's degrees of freedom, and hence, for higher lamba's the curve that is fit by our model is more ~~[illegible]~~ ~~[illegible]~~ rigid / more linear as compared to lower lambdas.

[Q-2.1]. Given $K(x, x')$ is valid kernel.

Then there has to exist $\phi : R^m \rightarrow H$ s.t

$$K(x, x') = \phi(x)^T \phi(x') \quad \text{--- (1)}$$

(i) Consider a function $\phi_{new} : R^m \rightarrow H$ s.t

$$\phi_{new}(x) = \phi(g(x)).$$

$$\therefore \phi_{new}(x)^T \phi_{new}(y) = \phi(g(x))^T \phi(g(y)).$$

Using eq^n-(1)

$$= K(g(x), g(y)).$$

Hence, $K(g(x), g(y))$ can be represented in terms of inner space product of some ϕ_{new}.

Hence $K_{new}(x, y) = K(g(x), g(y))$ is a valid kernel.

ii) Let $g(x) = \sum_{i=0}^{n} a_i x^i$

Property-1 : If K_1, K_2 are valid kernels, then $\alpha_1 K_1 + \alpha_2 K_2$ are valid kernels $\forall \alpha_1, \alpha_2 \geq 0$

[Proved in class]

Property-2. If K_1 & K_2 are valid kernels, then $K_1 \cdot K_2$ is also a valid kernel

(proved in class).

Property 3 :- If K_1 is a valid kernel, then $(K_1)^d$ is a valid kernel $\forall d \in N$.

(Proved in part 1).

Using property 1, 3 we can say that for ~~any i ∈ {0,1,...~~ any $i \in \{0, 1, \dots n-1\}$.

$a_i K(x, x')^i$ is a valid kernel.

Hence summation over all $i \in N$, $i < n$

$\sum_{i=0}^{i<n} a_i K(x, x')^i$ is also a valid kernel

Hence proved. $q(K(x, x'))$ is also a valid kernel

[Q. 2.2]

My kernel = $K(x,y) = (1 + x^T y)^4$

This gives an error < 7000 which is desired. [Plot attached on the next page]