Q.5 For any x_i, if we use linear approximation then we will be using only the first k columns of V which correspond to largest k eigenvalues. Hence, this approximation must yield the lowest reconstruction error for all reconstructions that use same k columns of V.

But, in case of non-linear approximation, we can choose different sets of k columns for different x_i.

eg $x_i \rightarrow \alpha_i^{(k)} = V_k^t x_i$

Say $\alpha_i = [1.5 \quad 0.7 \quad 0.3 \quad 1.8 \quad 2]$

then we can choose columns 1, 5 and 4 i.e. the highest values of α_i in case of non-linear approximation.

This will give lesser error as compared to choosing columns 1, 2, 3 in linear approximation

Algorithm :

$\alpha_i = V^T x_i$

β_i = zeros (size(α_i))

For indices with top k α_i :

$\beta_i = 1$

Time complexity = $O(D^3 + D) = O(D^2)$

This is a greedy approach.

$$\|x_i - V\beta_i\|^2 = [\|V^T\| \|x_i - V\beta_i\|]^2$$

$$= [\|V^T x_i - V^T V \beta_i\|]^2$$

$$= \|\alpha_i - \beta_i\|^2 \quad [\because V^T V = I \text{ orthogonal}]$$

Thus thus gives least error.

Since $\beta_i = 1$ where α_i's are largest