Ans-5)

A-5) The second option $E_L(V) \geq E_N(V)$ is correct.

In linear approximation, we have to choose first k columns of V corresponding to k largest eigenvalues.

In non-linear approximation, there is no restricting on choosing k columns. We can choose different k columns of V for different points, which minimizes reconstruction error.

Thus, the flexibility to choose k columns in non-linear approximation, makes second option $(E_L(V) \geq E_N(V))$ true.

Algorithm and Arguement why it is correct.

We'll calculate $\alpha_i = V^T x_i$ firstly. And, then choose $\alpha_i^{(k)} \rightarrow k$ entries from α_i with largest absolute value and set remaining to 0.

$$\underset{\alpha_i^{(k)}}{\operatorname{argmin}} \| x_i - V\alpha_i^{(k)} \|^2$$

~~$(\because V$ is orthonormal$)$~~

$$= \underset{\alpha_i^{(k)}}{\operatorname{argmin}} \| \alpha_i - \alpha_i^{(k)} \|^2 \quad (\because V \text{ is orthonormal})$$

which is minimized only with highest $\alpha_i^{(k)}$ values.

This is because the problem is to reduce reconstruction error which is reduced by highest values of α_i, since they have a higher contribution in constructing image.