Q1) $\hat{x} = VWx$

$$\Rightarrow L = \sum \|x - \hat{x}\|^2$$
$$= (X - VWX)^T (X - VWX)$$
$$= (X^T - X^T W^T V^T)(X - VWX)$$
$$= X^T X - X^T VWX - X^T W^T V^T X + X^T W^T V^T VWX$$
$$= X^T X - 2X^T VWX + (VWX)^2$$

now we want to minimize L

$$\Rightarrow \frac{dL}{dV} = 0 \quad \& \quad \frac{dL}{dW} = 0$$

$$\Rightarrow -2W^T (X - WVX) X^T = 0$$
$$\Rightarrow WV = I \quad \text{(assuming } X \text{ to be invertible)}$$

assuming W to be orthonormal (ie. $W^{-1} = W^T$)

$$\Rightarrow \boxed{V = W^T}$$

Q2) the covariance matrix: $\Sigma = \frac{1}{N} X^T X$

$$\Rightarrow L = N \cdot tr((I - W^T W) \Sigma (I - W^T W)^T)$$
$$= \underbrace{N\, tr(\Sigma)}_{\text{const.}} - N\, tr(\Sigma W^T W)$$

$$\Rightarrow L = -tr(\Sigma W^T W)$$

Now use the fact that Σ can be decomposed as:

$$\Sigma = Q \Lambda Q^T \qquad \left(\begin{array}{l} \Lambda = \text{eigen value matrix} \\ Q = \text{eigen vector matrix} \end{array} \right)$$

$$\Rightarrow tr(\Sigma W^T W) = tr(Q \Lambda Q^T W^T W)$$

$$= tr(\Lambda \underbrace{Q^T W^T W Q}_{U}) \qquad (\text{as } tr(AB) = tr(BA))$$

$$= tr(\Lambda U)$$

$$= \sum \lambda_i U_{ii}$$

$\Rightarrow$ to minimize L, we need to maximize $\sum \lambda_i U_{ii}$

$\therefore$ we need $U_{ii} = 1$ for indices corresponding to the k largest eigen values

$\Rightarrow$ W^* (optimal w) has rows forming an orthonormal basis for the subspace spanned by top k eigen vectors of Σ //