MSiA 421 HW01 Q4

4.) Reconstruction Error $= \frac{1}{N}\sum_{n=1}^{N} \|x_n - \hat{x}_n\|^2$

$$x_n = \sum_{i=1}^{D}(\alpha_{ni} u_i) \Rightarrow \alpha_{ni} = x_n^T u_i,\; x_n = \sum_{i=1}^{D}(x_n^T u_i) u_i$$

$$\hat{x}_n = \sum_{i=1}^{M}(x_n^T u_i) u_i + \sum_{i=M+1}^{D} b_i w_i = \sum_{i=1}^{M} z_{ni} u_i + \sum_{i=M+1}^{D} b_i u_i$$

$$\Rightarrow z_{ni} = x_n^T u_i$$

$$\Rightarrow x_n - \hat{x}_n = \sum_{i=1}^{D}(x_n^T u_i)u_i - \sum_{i=1}^{M}(x_n^T u_i)u_i - \sum_{i=M+1}^{D} b_i w_i$$

$$\Rightarrow z_{ni} = x_n^T u_i \Rightarrow \sum_{i=M+1}^{D} z_{ni} u_i - b_i u_i$$

$$\Rightarrow \frac{1}{N}\sum_{n=1}^{N} \left\| \sum_{i=M+1}^{D} z_{ni} u_i - b_i u_i \right\|^2$$

Taking derivatives with respect to b_j and z_{nj} and set $= 0$

$$\Rightarrow \frac{\partial RE}{\partial b_j} \Rightarrow b_j = \bar{x}^T u_j,\quad \frac{\partial RE}{\partial z_{nj}} \Rightarrow z_{nj} = x_n^T u_j$$

$$\Rightarrow \frac{1}{N}\sum_{n=1}^{N} \left\| \sum_{i=M+1}^{D} (x_n^T u_i - \bar{x}^T u_i) u_i \right\|^2 = \frac{1}{N}\sum_{n=1}^{N} \left\| \sum_{i=M+1}^{D} ((x_n - \bar{x})^T u_i) u_i \right\|^2$$

$$= \frac{1}{N}\sum_{n=1}^{N} \left(\sum_{i=M+1}^{D} ((x_n - \bar{x})^T u_i) u_i \right)^T \left(\sum_{j=M+1}^{D} ((x_n - \bar{x})^T u_j) u_j \right)$$

$$= \frac{1}{N}\sum_{n=1}^{N} \sum_{i=M+1}^{D} \sum_{j=M+1}^{D} ((x_n - \bar{x})^T u_i) u_i^T u_j ((x_n - \bar{x})^T u_j)$$

$\Rightarrow$

$$\Rightarrow \frac{1}{N} \sum_{n=1}^{N} \sum_{i=M+1}^{D} (x_n - \bar{x})^T u_i \, (x_n - \bar{x})^T u_i$$

$$= \frac{1}{N} \sum_{n=1}^{N} \sum_{i=M+1}^{D} u_i^T (x_n - \bar{x}) (x_n - \bar{x})^T u_i$$

$$= \sum_{i=M+1}^{D} u_i^T S \, u_i$$

$\Rightarrow$ Since orthonormal basis $\Rightarrow \|u\|^2 = 1$. This is so $w = 0$ can't occur.

$\Rightarrow$ Example: $D = 10$, $M = 1$

$\Rightarrow$ minimize $RE = u_{10}^T S u_{10} + \lambda_{10} (1 - u_{10}^T u_{10})$

$\Rightarrow \frac{\partial RE}{\partial u_{10}}$ and set $= 0 \Rightarrow 0 = u_{10}^T S - u_{10}^T \lambda_{10}$

$\Rightarrow S u_{10} = \lambda_{10} u_{10} \Rightarrow$ more generally $\Rightarrow S u_i = \lambda_i u_i$

$\Rightarrow$ Minimum R.E. $= \sum_{i=M+1}^{D} \lambda_i$