Benjamin Lapostolle

Question 1

$j_1, j_2 \in [\![1, m]\!]$, the sample covariance of $C_i v_{j_1}^*, C_i v_{j_2}^*$ is :

$$\mathrm{Cov}(C_i v_{j_1}^*, C_i v_{j_2}^*) = \frac{1}{n-1} \sum_{k=1}^{n} \left([C_i v_{j_1}^*]_{k1} - \mathbb{E}(C_i v_{j_1}^*)\right) \times \left([C_i v_{j_2}^*]_{k1} - \mathbb{E}(C_i v_{j_2}^*)\right)$$

$j \in [\![1, m]\!]$,

$$\mathbb{E}(C_i v_j^*) = \frac{1}{n} \sum_{k=1}^{n} [C_i v_j^*]_{k1}$$
$$= \frac{1}{n} \sum_{k=1}^{n} \sum_{l=1}^{m} [C_i]_{kl} [v_j^*]_{l1}$$
$$= \frac{1}{n} \sum_{l=1}^{m} [v_j^*]_{l1} \underbrace{\sum_{k=1}^{n} [C_i]_{kl}}_{=0}$$

$= 0$ because $C_i = X_i - M_i$, hence, $\mathbb{E}(C_i[:, l]) = 0$ $l \in [\![1, m]\!]$,

Therefore,

$$\mathrm{Cov}(C_i v_{j_1}^*, C_i v_{j_2}^*) = \frac{1}{n-1} \sum_{k=1}^{n} [C_i v_{j_1}^*]_{k1} [C_i v_{j_2}^*]_{k1}$$
$$= \frac{1}{n-1} [C_i v_{j_1}^*]^T C_i v_{j_2}^*$$
$$= \frac{1}{n-1} v_{j_1}^{*T} C_i^T C_i v_{j_2}^*$$

$k, l \in [\![1, m]\!]$,

$$[C_i^T C_i]_{kl} = \sum_{k_2=1}^{n} [C_i]_{k_2 k} [C_i]_{k_2 l}$$

$$(\mathrm{Var}(C_i))_{kl} = \frac{1}{n-1}\sum_{k=1}^{m}([C_i]_{bk} - 0)([C_i]_{bl} - 0)$$

$$= \frac{1}{n-1}[C_i^T C]_{kl}$$

Hence: $j_1, j_2 \leq m$,

$$\mathrm{Cov}(C_i U_{j_1}^*, C_i U_{j_2}^*) = \cancel{\frac{1}{(n-1)^2}} U_{j_1}^{*T} \mathrm{Var}(C_i) U_{j_2}^*$$

$$= \cancel{\frac{1}{(n-1)^2}} \times \lambda_{j_2} U_{j_1}^{*T} U_{j_2}^*$$

λ_{j_2} eigenvalue of $U_{j_2}^*$

$$= \begin{cases} 0 & \text{if } j_1 \neq j_2 \text{ (orthogonal vector basis)} \\ \frac{\lambda_{j_1}}{\cancel{(n-1)^2}} & \text{else} \end{cases}$$

Therefore, we have that the true principal components $C_i U_j^*, j \in \{1, \dots, m\}$ are uncorrelated;

and: $\boxed{j \in \{1, \dots m\}, \mathrm{Var}(C_i U_j^*) = \frac{\lambda_j}{\cancel{(n-1)^2}}}$

Question 2

i) The quality of the images increases with k. When k is low ($k \leq 10$), it seems that the quality increases highly, whereas for $30 \leq k \leq 60$, the gain in quality with k seems to ~~be~~ decrease (the overall quality still increases of course).

ii) PCA and MDS methods seem to be the methods producing the least lossy compression of the image.

Question 3

$$C_i' = X_i' - M_i' = X_i Q - M_i'$$

$$j \le m, \quad [M_i']_{\cdot j} = \mathbb{E}\left([X_i'][:,j]\right)$$

$$= \frac{1}{n}\sum_{k=1}^{n} [X_i']_{kj}$$

$$= \frac{1}{n}\sum_{k=1}^{n} [X_iQ]_{kj}$$

$$= \frac{1}{n}\sum_{k=1}^{n}\sum_{l=1}^{m} [X_i]_{kl}[Q]_{lj}$$

$$= \frac{1}{n}\sum_{l=1}^{m} [Q]_{lj}\sum_{k=1}^{n} [X_i]_{kl}$$

$$= \cancel{\frac{1}{n}}\sum_{l=1}^{m} [Q]_{lj} \times [M_i]_{\cdot l}$$

$$= \sum_{l=1}^{m} [M_i]_{\cdot l}[Q]_{lj}$$

$$[M_i']_j = [M_iQ]_{\cdot j}$$

Hence: $M_i' = M_iQ$

So: $C_i' = X_iQ - M_iQ = C_iQ$

$$i,j \in [1,m] \quad (\mathrm{Var}(C_i'))_{ij} = (\mathrm{Var}(C_iQ))_{ij} = \frac{1}{n-1}\sum_{k=1}^{n} [C_iQ]_{ki}[C_iQ]_{kj}$$

$$= \frac{1}{n-1}\left[(C_iQ)^T C_iQ\right]_{ij}$$

$$\Rightarrow \mathrm{Var}(C_i') = \frac{1}{n-1}(C_iQ)^T C_iQ = \frac{1}{n-1}Q^T C_i^T C_i Q$$

$$= \frac{1}{n-1}Q^T (n-1)\mathrm{Var}(C_i) Q$$

$$= Q^T \mathrm{Var}(C_i) Q$$

Hence, with $Var(C_i) = O^T diag(\lambda_1, \dots, \lambda_m) O$

$Var(C_i') = Q^T O^T diag(\lambda_1, \dots, \lambda_m) Q O$

$Var(C_i')$ and $Var(C_i)$ have the same eigenvalues. Therefore the low rank approximations capture equal amounts of the variance of the two images, because the eigenvalues are linked to the variance.

$j \leq m$ • Is $C_i U_j = C_i' U_j'$?

$C_i' U_j' = C_i Q U_j'$. Let's show that QU_j' is an eigenvector of eigenvalue λ_j of $Var(C_i)$.

$$\begin{aligned} I_m Var(C_i) Q U_j' &= Q Q^T Var(C_i) Q U_j' \\ &= Q Q^T O diag(\lambda_1, \dots, \lambda_m) O Q U_j' \\ &= Q Var(C_i') U_j' \\ Var(C_i) Q U_j' &= \lambda_j Q U_j' \end{aligned}$$

Hence, QU_j' is an eigenvector of eigenvalue λ_j. Therefore, since U_j is also an eigenvector of eigenvalue λ_j, we have that:

$QU_j' = U_j$ (assuming they have the same direction ie sign).

Therefore:

$j \leq m$, $\boxed{C_i U_j = C_i Q U_j' = C_i' U_j'}$

Question 4

The low rank approximation allows us to remove most of the noise from the image. In the case of the second logo, few noises remain. However for the first image, even if ~~the~~ noises are removed, the logo is quite blur so it did not work as well.

Question 5

When we use a PCA method, we choose $k < m$ directions, and we keep only the projection of the data on these directions. Therefore, if the information corresponding to the noise in the image is located in directions such that their projections ~~is~~ are little on the k chosen directions, the noise will be removed. That must be the case for the second logo.

However, if the information of the noise is located in such a way that by projecting the data on the k directions we still have most of the information, then the noise won't disappear.
That is why we find $k=4$ for the first logo, because the $k \geq 5$ directions must getting too much information of the noise.