# STAT154 HW5
# PCA and clustering

Seth Metcalf

7 April 2024

## 1 Theoretical Exercises

### Q1 (Principal component analysis, formulation 1)

#### (a)

Given

$$\arg\max_{\|\boldsymbol{\phi}_1\|_2^2=1} \frac{1}{n}\sum_{i=1}^{n}\langle \boldsymbol{x}_i, \boldsymbol{\phi}_1\rangle^2 = \arg\max_{\|\boldsymbol{\phi}_1\|_2^2=1}\langle \boldsymbol{\phi}_1, \boldsymbol{X}^{\mathsf{T}}\boldsymbol{X}\boldsymbol{\phi}_1\rangle.$$

, we can then use this to show:

Compute the expansion of $\langle v, X^T X v\rangle$:

$$\langle v, X^T X v\rangle = v^T X^T X v$$

$$= v^T V \Sigma^T U^T U \Sigma V^T v$$

$$= v^T V \Sigma^T I \Sigma V^T v$$

$$= v^T V \Sigma^T \Sigma V^T v$$

$$= \|\Sigma V^T v\|^2$$

$$= \|[\sigma_1 \ 0 ... 0]^T \begin{bmatrix} v_1^T \\ \vdots \\ v_d^T \end{bmatrix} \|^2$$

$$= \sigma_1^2 \| \begin{bmatrix} v_1^T \\ \vdots \\ v_d^T \end{bmatrix} \|^2$$

$$= \sigma_1^2 \|v\|^2$$

$$= \sigma_1^2$$

Since it is given that $\sigma_1 > 0$, it follows $\langle v, X^T X v\rangle < \sigma_1^2$

Noted above, it is shown that $\{v_1, -v_1\}$ achieves the maximum value of $\langle \boldsymbol{\phi}_1, X^T X \boldsymbol{\phi}_1\rangle$, any vector other than $v_1$ and $-v_1$ is shown (seen above) to be less than $\sigma_1$, meaning that the optimal value/argmax ($\boldsymbol{\phi}_1^*$) is $\{v_1, -v_1\}$

**(b)**

Prove that $\langle v, X^T X v\rangle < \sigma_2^2$, since we are now using $v_2$ we can compute that $v_\perp$ is orthogonal to $v_1$, which follows that:

$$\begin{aligned}\langle v, X^T X v\rangle &= \langle v_\perp + cv_1, X^T X(v_\perp + cv_1)\rangle \\ &= \langle v_\perp, X^T X v_\perp\rangle + c^2\langle v_1, X^T X v_1\rangle\end{aligned}$$

And since $\sigma_1 > \sigma_2 > 0$, we know that for any $c$, $c^2\sigma_1^2$ will always be less than $\sigma_2^2$, which then follows that $\langle v, X^T X v\rangle < \sigma_2^2$. Because of this, similar reasoning to before, $\{v_2, -v_2\}$ achieves the maximum value of $\langle \boldsymbol{\phi}_2, X^T X \boldsymbol{\phi}_2\rangle$ and any other vector orthogonal to $v_1$ but not equal to $v_2, -v_2$ yields a value less than $\sigma_2^2$, meaning the argmax is $\{v_2, -v_2\}$.

## Q2 (Principal component analysis, formulation 2)

**(a)**

Show that for any $\boldsymbol{x} \in \mathbb{R}^d$,

$$\|\boldsymbol{x} - \mathsf{P}_S\boldsymbol{x}\|_2^2 = \min_{z_1,\dots z_M \in \mathbb{R}} \left\|\boldsymbol{x} - \sum_{k=1}^{M} z_k \boldsymbol{\phi}_k\right\|_2^2,$$

Which means that the error $\|\boldsymbol{x} - \mathsf{P}_S\boldsymbol{x}\|_2^2$ is minimized when $\boldsymbol{x}$ ius approximated by its projection onto subspace $S$, spanned by the orthonormal vectors $\{\boldsymbol{\phi}_1, ..., \boldsymbol{\phi}_M\}$.

Projection of $\boldsymbol{x}$ onto $S$ is given by:

$$\mathsf{P}_S\boldsymbol{x} = \sum_{k=1}^{M} (\boldsymbol{x} \cdot \boldsymbol{\phi}_k)\boldsymbol{\phi}_k$$

Therefore:

$$\|\boldsymbol{x} - \mathsf{P}_S\boldsymbol{x}\|_2^2 = \|\boldsymbol{x} - \sum_{k=1}^{M} (\boldsymbol{x} \cdot \boldsymbol{\phi}_k)\boldsymbol{\phi}_k\|_2^2 = \|\boldsymbol{x} - \sum_{k=1}^{M} z_k \boldsymbol{\phi}_k\Big\|_2^2$$

Meaning that when $z_k = \boldsymbol{x} \cdot \boldsymbol{\phi}_k$, $\boldsymbol{x} \cdot \boldsymbol{\phi}_k$ showing the projection of $\boldsymbol{x}$ onto the subspace $S$. To expand on this idea, we can then show that:

$$\sum_{i=1}^{n} \|\boldsymbol{x}_i - \mathsf{P}_S\boldsymbol{x}_i\|_2^2 = \min_{(z_{ik})_{i\in[n],k\in[M]}} \sum_{i=1}^{n} \left\|\boldsymbol{x}_i - \sum_{k=1}^{M} z_{ik}\boldsymbol{\phi}_k\right\|_2^2,$$

Because the error for each $\boldsymbol{x}_i$ is minimized when it is approximated by its projection/expressed as a linear combination of the basis vectors. Therefore, the total error is minimized when each $\boldsymbol{x}_i$ is expressed in terms of its projection.

**(b)**

In order to show that

$$\min_{S\in S_M} \sum_{i=1}^{n} \|\boldsymbol{x}_i - P_S\boldsymbol{x}_i\|_2^2 = \min_{\boldsymbol{V}_M \in M} \min_{\boldsymbol{Z}\in\mathbb{R}^{n\times M}} \|\boldsymbol{X} - \boldsymbol{Z}\boldsymbol{V}_M^{\mathsf{T}}\|_F^2,$$

The error term $\|\boldsymbol{x}_i - \mathsf{P}_S\boldsymbol{x}_i\|_2^2$, is minimized in a given subspace $S$ for any $\boldsymbol{x}_i$ by approximating its projection. The optimal subspace is given by the first $M$ principal components, $\boldsymbol{\phi}_1, ..., \boldsymbol{\phi}_M$, which is equivalent to $v_1, ..., v_M$ which, in this case, is given by $V_M$.

If we consider $X$ as a matrix where each row represents $\boldsymbol{x}_i$ data point, we can express $X$ as $X = ZV_M^T$ where $Z$ is a matrix containing the coeff of each $\boldsymbol{x}_i$ along the principal components.

Therefore, we can see on the right side of the problem $\|\boldsymbol{X} - \boldsymbol{Z}\boldsymbol{V}_M^{\mathsf{T}}\|_F^2$ is the frobenius norm between the matrix $X$, and its approximation with the first $M$ principal components $(X - ZV_M^T)$.

Thus, the minimum error approximating $X$ by its projection is equivalent to minimizing the Frobenius norm of the matrix $X$ and its approximation with the first $M$ principal components as seen in the equation.

## Q3 (K-means algorithm with A-norm)

### (a)

Prove that

$$\mathrm{WCV}(C_k) \equiv \sum_{i,j \in C_k} \|\boldsymbol{x}_i - \bar{\boldsymbol{x}}_{C_k}\|_{\boldsymbol{A}}^2,$$

where

$$\bar{\boldsymbol{x}}_{C_k} = \frac{1}{|C_k|} \sum_{j \in C_k} \boldsymbol{x}_j.$$

Knowing this, we can look at the expansion of $\mathrm{WCV}(C_k)$:

$$\mathrm{WCV}(C_k) \equiv \frac{1}{2|C_k|} \sum_{i,j \in C_k} \|\boldsymbol{x}_i - \bar{\boldsymbol{x}}_{C_k}\|_{\boldsymbol{A}}^2$$

$$= \frac{1}{2|C_k|} \sum_{i,j \in C_k} \langle x_i - x_j, \boldsymbol{A}(x_i - x_j)\rangle$$

$$= \frac{1}{2|C_k|} \sum_{i,j \in C_k} (\langle x_i, \boldsymbol{A}\boldsymbol{x}_i\rangle - 2\langle x_i, \boldsymbol{A}\boldsymbol{x}_j\rangle + \langle x_j, \boldsymbol{A}\boldsymbol{x}_j\rangle)$$

$$= \sum_{i \in C_k} \langle x_i, \boldsymbol{A}\boldsymbol{x}_i\rangle - 2 \sum_{i,j \in C_k} \langle x_i, \boldsymbol{A}\boldsymbol{x}_j\rangle + \sum_{j \in C_k} \langle x_j, \boldsymbol{A}\boldsymbol{x}_j\rangle$$

Substitute in $\bar{\boldsymbol{x}}_{C_k}$:

$$= \sum_{i \in C_k} \langle x_i, \boldsymbol{A}\boldsymbol{x}_i\rangle - 2|C_k|\langle \bar{\boldsymbol{x}}_{C_k}, \boldsymbol{A}\bar{\boldsymbol{x}}_{C_k}\rangle + \sum_{j \in C_k} \langle x_j, \boldsymbol{A}\boldsymbol{x}_j\rangle$$

$$= \sum_{i \in C_k} \langle x_i, \boldsymbol{A}\boldsymbol{x}_i\rangle - 2|C_k|\|\bar{\boldsymbol{x}}_{C_k}\|_A^2 + \sum_{j \in C_k} \langle x_j, \boldsymbol{A}\boldsymbol{x}_j\rangle$$

$$= \sum_{i \in C_k} (\langle x_i, \boldsymbol{A}\boldsymbol{x}_i\rangle - 2\|\bar{\boldsymbol{x}}_{C_k}\|_A^2 + \langle x_j, \boldsymbol{A}\boldsymbol{x}_j\rangle)$$

$$= \sum_{i \in C_k} (\langle x_i, \boldsymbol{A}\boldsymbol{x}_i\rangle - 2\langle \bar{x}_{C_k}, \boldsymbol{A}\bar{\boldsymbol{x}}_{C_k}\rangle + \langle x_j, \boldsymbol{A}\boldsymbol{x}_j\rangle)$$

$$= \sum_{i \in C_k} (\|\boldsymbol{x}_i\|_A^2 - 2\langle \bar{x}_{C_k}, \boldsymbol{A}\bar{\boldsymbol{x}}_{C_k}\rangle + \|\bar{\boldsymbol{x}}_{C_k}\|_A^2)$$

$$= \sum_{i \in C_k} (\|\boldsymbol{x}_i - \bar{\boldsymbol{x}}_{C_k}\|_A^2)$$

Therefore,

$$\mathrm{WCV}(C_k) \equiv \sum_{i \in C_k} (\|\boldsymbol{x}_i - \bar{\boldsymbol{x}}_{C_k}\|_A^2)$$

### (b)

Since we know that

$$R((C_k)_{k\in[K]}) = \sum_{k=1}^{K} \text{WCV}(C_k).$$

We can rewrite this as:

$$R((C_k)_{k\in[K]}) = \sum_{k=1}^{K} \sum_{i\in C_k} (\|\boldsymbol{x}_i - \bar{\boldsymbol{x}}_{C_k}\|_A^2)$$

Also, we know that

$$\overline{R}((w_{ik})_{i\in[n],k\in[K]}, (\boldsymbol{\mu}_k)_{k\in[K]}) = \sum_{i=1}^{n} \sum_{k=1}^{K} \|\boldsymbol{x}_i - \boldsymbol{\mu}_k\|_{\boldsymbol{A}}^2 w_{ik}.$$

To prove

$$\min_{(C_k)_{k\in[K]}} R((C_k)_{k\in[K]}) = \min_{(w_{ik})_{i\in[n],k\in[K]}\in\mathcal{W}} \min_{(\boldsymbol{\mu}_k)_{k\in[K]}} \overline{R}((w_{ik})_{i\in[n],k\in[K]}, (\boldsymbol{\mu}_k)_{k\in[K]}), \tag{1}$$

We need to show that left-hand side is less than or equal to the right-hand side and the right-hand side is less than or equal to the left-hand side.

Left hand side $\leq$ Right hand side: Given a partition $(C_k)_{k\in[K]}$ we can compute the centroids $\mu_k$ and the weights $w_{ik}$ according to the assignment of clusters, so weights and centroids can always be constructed. This further implies that the minimum over all partitions is $\leq$ than the minimum over all possible weights and centroids from the set, meaning that we have:

$$\min_{(C_k)_{k\in[K]}} R((C_k)_{k\in[K]}) \leq \min_{(w_{ik})_{i\in[n],k\in[K]}\in\mathcal{W}} \min_{(\boldsymbol{\mu}_k)_{k\in[K]}} \overline{R}((w_{ik})_{i\in[n],k\in[K]}, (\boldsymbol{\mu}_k)_{k\in[K]})$$

Right hand side $\leq$ left hand side: We have to show that any assignment of weights and centroids that minimizes $\overline{R}$ there is a partition $(C_k)_{k\in[K]}$ that minimizes $R$. Since we can assign each data point $\boldsymbol{x}_i$ to cluster $k$ that has minimum distance by the weighted $A - norm$, the partition creates has to minimize $R$. Which shows us that:

$$\min_{(w_{ik})_{i\in[n],k\in[K]}\in\mathcal{W}} \min_{(\boldsymbol{\mu}_k)_{k\in[K]}} \overline{R}((w_{ik})_{i\in[n],k\in[K]}, (\boldsymbol{\mu}_k)_{k\in[K]}) \leq \min_{(C_k)_{k\in[K]}} R((C_k)_{k\in[K]})$$

Since both parts have been proven with help from part (a) then we have shown that:

$$\min_{(C_k)_{k\in[K]}} R((C_k)_{k\in[K]}) = \min_{(w_{ik})_{i\in[n],k\in[K]}\in\mathcal{W}} \min_{(\boldsymbol{\mu}_k)_{k\in[K]}} \overline{R}((w_{ik})_{i\in[n],k\in[K]}, (\boldsymbol{\mu}_k)_{k\in[K]})$$

### (c)

Have to minimize $\overline{R}$ with respect to $(\mu_k)_{k\in[K]}$, we differentiate $\overline{R}$ with respect to $\mu_k$ and set deriv = 0:

$$\frac{\partial \overline{R}}{\partial \mu k} = \frac{\partial}{\partial \mu_k}(\sum_{i=1}^{n} \sum_{k=1}^{K} \|\boldsymbol{x}_i - \mu_k\|_A^2 w_{ik})$$

$$= \frac{\partial}{\partial \mu_k}(\sum_{i=1}^{n} \|\boldsymbol{x}_i - \mu\|_A^2 w_{ik})$$

$$= -2 \sum_{i=1}^{n} \|\boldsymbol{x}_i - \mu\|_A^2 w_{ik} \boldsymbol{A}$$

$$= -2\boldsymbol{A}(\sum_{i=1}^{n} \boldsymbol{x}_i w_{ik} - \mu \sum_{i=1}^{n} w_{ik})$$

$$= -2\boldsymbol{A}(\sum_{i=1}^{n} \boldsymbol{x}_i w_{ik} - \mu_k n_k)$$

Now, we set the derivative to 0:

$$-2\boldsymbol{A}(\sum_{i=1}^{n} \boldsymbol{x}_i w_{ik} - \mu_k n_k) = 0$$

$$\sum_{i=1}^{n} \boldsymbol{x}_i w_{ik} - \mu_k n_k = 0$$

$$\mu_k n_k = \sum_{i=1}^{n} \boldsymbol{x}_i w_{ik}$$

$$\mu_k = \frac{1}{n_k} \sum_{i=1}^{n} \boldsymbol{x}_i w_{ik}$$

Therefore, the minimizer of $\mu_k$ is given by the centroids calculated using the weighted mean of the data points for each cluster.

### (d)

Same idea as (c), we differentiate $\overline{R}$ with respect to $w_{ik}$ and set deriv = 0:

$$\frac{\partial \overline{R}}{\partial w_{ik}} = \frac{\partial}{\partial w_{ik}}(\sum_{i=1}^{n} \sum_{k=1}^{K} \|\boldsymbol{x}_i - w_{ik}\|_A^2 \mu_{ik})$$

$$\vdots$$

$$w_{ik} = \frac{1}{n_k} \sum_{i=1}^{n} \boldsymbol{x}_i \mu_k$$

# 2 Computational Exercises

## Q1 (PCA on the Olivetti faces dataset)

Source code:

```
# Computational Exercises Q1
import numpy as np
import matplotlib.pyplot as plt
from sklearn.datasets import fetch_olivetti_faces
from sklearn.decomposition import PCA

# Load Olivetti faces
data = fetch_olivetti_faces()
X = data.data
n_samples, n_features = X.shape
```

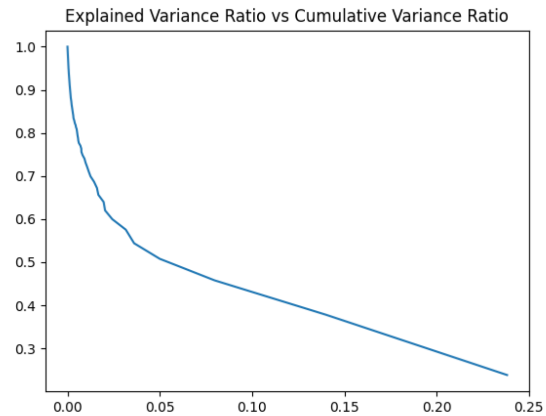


**Figure 1:** Q1 Olvetti Dataset

```
# Center
X_centered = X - np.mean(X, axis=0)

# PCA
pca = PCA()
pca.fit(X_centered)

ex_var_ratio = pca.explained_variance_ratio_
cum_var_ratio = np.cumsum(ex_var_ratio)

#print(ex_var_ratio)
#print(cum_var_ratio)

# Plot
plt.plot(ex_var_ratio, cum_var_ratio)
plt.title('Explained Variance Ratio vs Cumulative Variance Ratio')
plt.show()
```

When $j > 400$ the proportion of variance explained becomes zero, this is because there are only 400 samples in the Olivetti faces dataset so the maximum number of principal components is also 400.