Projet
Application de l'ACP : les Eigenfaces Méthode de la puissance V2

Najmeddine Ayoub

Département Sciences du Numérique
2020-2021

Table des matières

1 Introduction

Ce document représente le rapport de la première partie du projet de calcul scientifique et analyse de donnée.

Cette première partie s'intéresse à l'étude de l'ACP (L'analyse en composante principale). Cet outil classique et puissant, permet l'étude des échantillons d'énormes tailles caractérisées par un nombre trés élevés de variables , en les visualisant sur des espaces de faibles di- mensions. La méthode essaye aussi de récupérer un pourcentage extrémement faiable de l'information apportée par l'échantillon étudié en se basant sur des outils de l'algébre linéaire.

L'ACP permet aussi une classification claire des classes contenues dans l'échantillon et facilite le partitionnement des données en clusters et cette étude peut étre menée grâce à des outils algorithmiques qui différent en terme performance et de complexité.

2 Les Eigenface

2.1 Question 1

Le calcule des axes principaux des images d'apprentissage nous donne les figures suivantes :

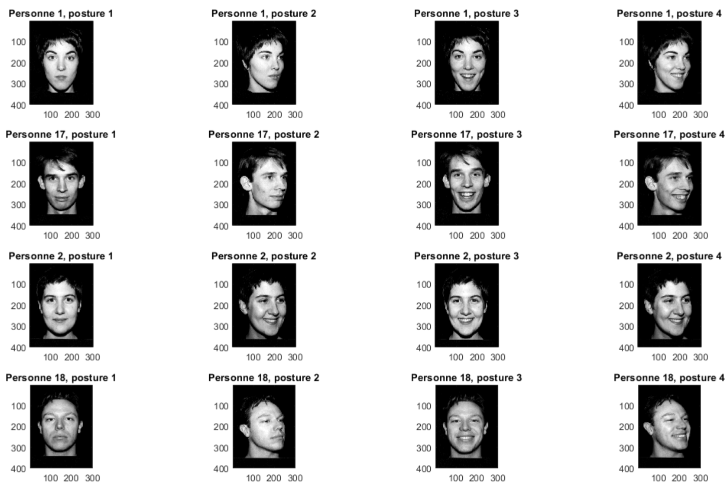

FIGURE 1 – Individus.

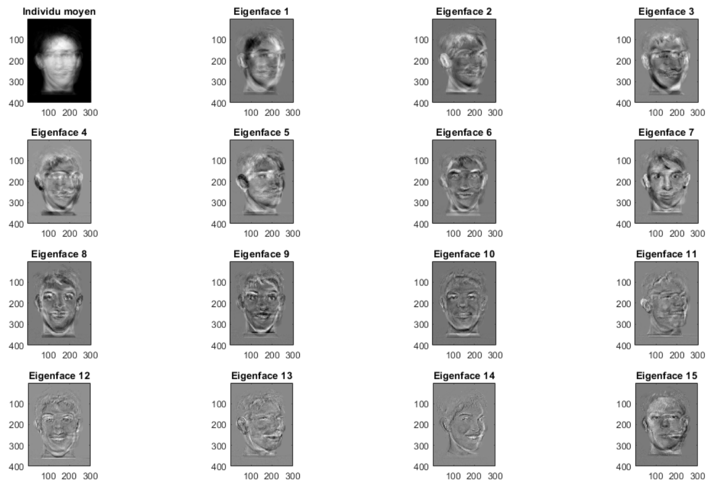

FIGURE 2 – Eigenfaces correspondants aux individus.

2.2 Question 2

Dans cette question on s'intéresse à l'affichage des es images d'apprentissage reconstruites à l'aide des q premières eigenfaces et des q premères composantes principales, pour $q \in [0, n-1]$. le script *projection.m* donne les images suivantes.

FIGURE 3 – images d'invidus reconstruites.

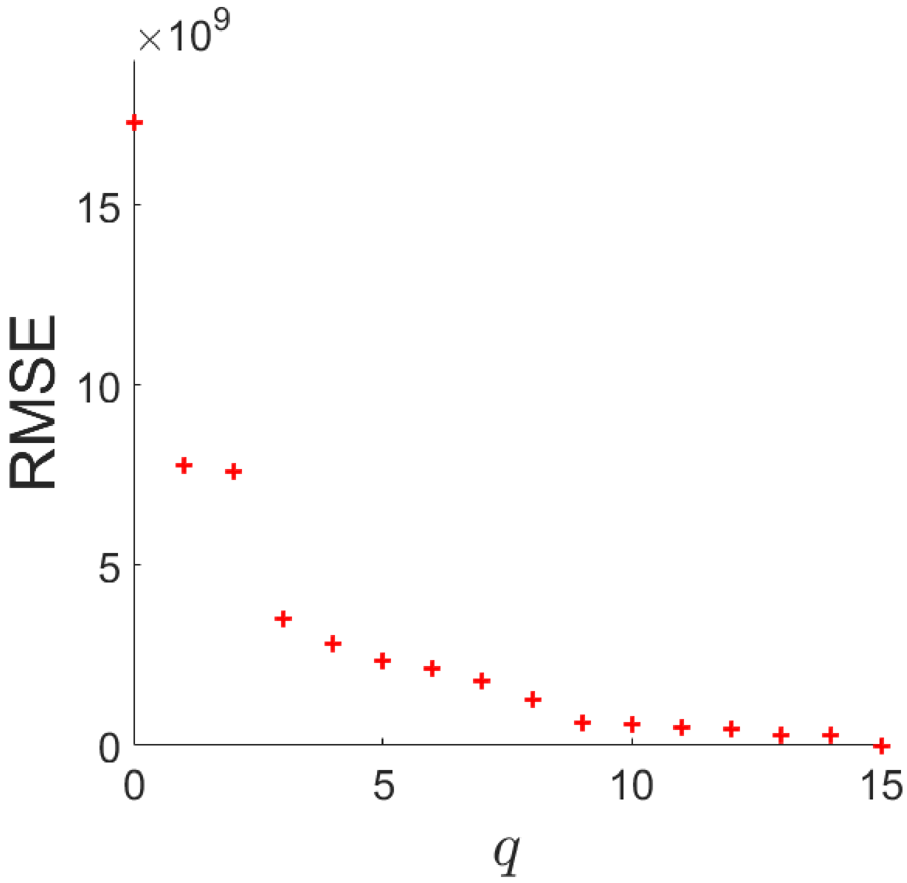

FIGURE 4 – Eigenfaces correspondants aux individus.

On remarque que le RMSE entre les images originales et les images ainsi reconstruites décroit en fonction de q, cela est due au fait que l'information est concentré sur les premieres composantes principales. plus q croit plus la contribution de ces composantes augmente dans la reconstitution de nos images.

2.3 Question 3

On produit des eingenfaces comme précedement pour des visages masqués, le masque est assimilé à un rectangle en noir qui cache la partie inférieur du visage, et obtient les images suivantes.

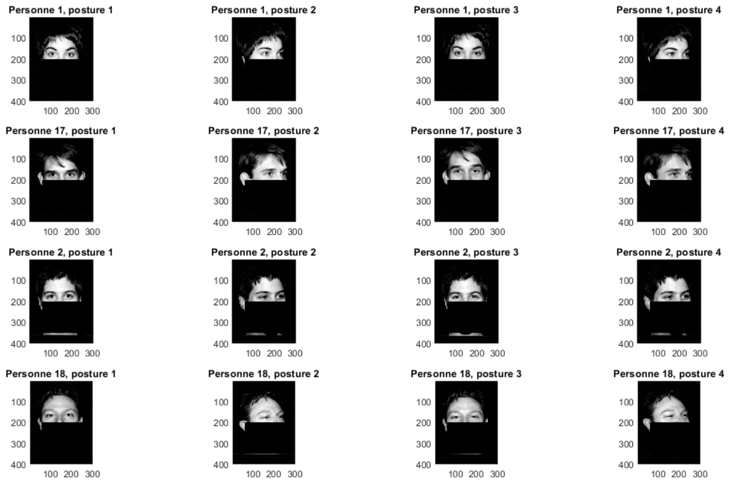

FIGURE 5 – Eigenfaces correspondants aux individus.

3 L'ACP et la méthode de la puissance itérée

3.1 Question 4

On a, pour tout $X \in M_{n,1}$:

$$X \in ker(H^tH - \lambda I_n) \Leftrightarrow H^tHX = \lambda X \Leftrightarrow \begin{cases} H^tY = \lambda X \\ Y = HX \end{cases} \Rightarrow \begin{cases} HH^tY = \lambda HX \\ Y = HX \end{cases} \Rightarrow Y \in ker(HH^t - \lambda I_n)$$

Donc la connaisance des éléments propres de H^tH permet de connaitre les éléments propres de HH^t.

3.2 Question 5

Le script *puissanceiteree.m* donne une méthode classique du calcul des couple propres dominants d'une matrice. On applique cette méthode à nos deux matrices H^tH et HH^t. Le calcul de l'erreur relative de cette méthodes pour les deux matrices, ainsi que le le temps d'exécution de la méthode donne le résultat suivant :

```
Erreur relative pour la methode avec la grande matrice = 9.887e-09
Erreur relative pour la methode avec la petite matrice = 9.787e-09
Ecart relatif entre les deux valeurs propres trouvees = 0.00e+00
Temps pour une ite avec la grande matrice = 7.977e-03
Temps pour une ite avec la petite matrice = 2.729e-04

Valeur propre dominante (methode avec la grande matrice) = 9.326e+04
Valeur propre dominante (methode avec la petite matrice) = 9.326e+04
Valeur propre dominante (fonction eig) = 9.326e+04
>>
```

FIGURE 6 – résultat d'exécution du script.

3.3 Question 6

Dans la méthode de la puissance iérée on divise par la valeur propre ce qui peut engendrer des erreurs graves car la matrice HH^t n'est pas inversible. Delà le choix de la fonction eig est le plus adapté à ce calcul.

3.4 Question 7

Pour calculer les éléments propre de Σ, on doit appliquer la méthode de la puissance itérée à la matrice de taille la plus petite, ç-à-d à la matrice AA^t de taille $p \times p$ avec (p<n).