

TP : Introduction à l'apprentissage statistique

## 1 Notations

Les données d'apprentissage sont les vecteurs  $(\mathbf{x}_i)_{i \in \llbracket 1, n \rrbracket}$  de  $\mathbb{R}^p$ . Leur moyenne empirique est notée  $\bar{\mathbf{x}}$ .

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \quad (1)$$

Et la matrice de covariance empirique  $\hat{\mathbf{R}}$ .

$$\hat{\mathbf{R}} = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \quad (2)$$

Cette matrice est semi-définie positive. On note  $(\lambda_i)_{i \in \llbracket 1, p \rrbracket}$  la suite croissante de ses valeurs propres, et  $(\mathbf{u}_i)_{i \in \llbracket 1, p \rrbracket}$  la suite des vecteurs propres orthonormés associés.

L'énergie de projection suivant la direction  $\mathbf{v}_j$  est définie par  $\mathbf{E}_j$ .

$$\mathbf{E}_j = \frac{1}{n} \sum_{i=1}^n \langle \mathbf{v}_j, \mathbf{x}_i - \bar{\mathbf{x}} \rangle^2 \quad (3)$$

Le ratio d'énergie de projection est défini par  $\mathcal{K}(l)$  où  $l$  est la dimension du sous-espace de projection  $\mathcal{S} = \text{vect}(\mathbf{u}_1, \dots, \mathbf{u}_l)$ .

$$\mathcal{K}(l) = \frac{\frac{1}{n} \sum_{i=1}^n \|\pi_{\mathcal{S}}(\mathbf{x}_i - \bar{\mathbf{x}})\|_2^2}{\frac{1}{n} \sum_{i=1}^n \|\mathbf{x}_i - \bar{\mathbf{x}}\|_2^2} \quad (4)$$

Les données sont regroupées dans une matrice  $\mathbf{X}$  de dimension  $(p, n)$ .

$$\mathbf{X} = \frac{1}{\sqrt{n}} [\mathbf{x}_1 - \bar{\mathbf{x}}, \dots, \mathbf{x}_n - \bar{\mathbf{x}}] \quad (5)$$

## 2 Principes de l'ACP

### 2.1 Direction maximisant l'énergie de projection

*Montrons que le vecteur propre  $\mathbf{u}_1$  de  $\hat{\mathbf{R}}$  maximise l'énergie de projection parmi tous les vecteurs unitaires.*

Notons  $\forall i \in \llbracket 1, n \rrbracket, \mathbf{y}_i = \mathbf{x}_i - \bar{\mathbf{x}}$ . On a :

$$\begin{aligned} \mathbf{E}_1 &= \frac{1}{n} \sum_{i=1}^n \langle \mathbf{v}_1, \mathbf{y}_i \rangle^2 \\ &= \frac{1}{n} \sum_{i=1}^n (\mathbf{v}_1 \mathbf{y}_i) (\mathbf{v}_1 \mathbf{y}_i)^T \\ &= \frac{1}{n} \sum_{i=1}^n \mathbf{v}_1 \mathbf{y}_i \mathbf{y}_i^T \mathbf{v}_1^T \\ &= \mathbf{v}_1 \left( \frac{1}{n} \sum_{i=1}^n \mathbf{y}_i \mathbf{y}_i^T \right) \mathbf{v}_1^T \\ &= \mathbf{v}_1 \hat{\mathbf{R}} \mathbf{v}_1^T \end{aligned}$$

On cherche à maximiser  $E_1$  en fonction du vecteur  $\mathbf{v}_1$  sous la contrainte  $\|\mathbf{v}_1\|_2 = 1$ . Ce problème peut se résoudre à l'aide d'un Lagrangien  $\mathcal{L}$ .

Soit  $\lambda \in \mathcal{R}$ , on pose  $\mathcal{L}(\mathbf{v}_1, \lambda) = \mathbf{E}_1(\mathbf{v}_1) - \lambda(\mathbf{v}_1 \mathbf{v}_1^T - 1)$ .

On dérive partiellement :

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \mathbf{v}_1}(\mathbf{v}_1, \lambda) &= 2\hat{\mathbf{R}}\mathbf{v}_1 - 2\lambda\mathbf{v}_1 \\ \frac{\partial \mathcal{L}}{\partial \lambda}(\mathbf{v}_1, \lambda) &= 1 - \mathbf{v}_1 \mathbf{v}_1^T \end{cases} \quad (6)$$

$$(7)$$

On pose (6) = 0 et (7) = 0, et on obtient :

$$\begin{cases} (\hat{\mathbf{R}} - \lambda I_n)\mathbf{v}_1 &= 0 \\ \|\mathbf{v}_1\|_2 &= 1 \end{cases} \quad (8)$$

$$(9)$$

Grâce à (8), on sait que  $\mathbf{v}_1$  est un vecteur propre associé à une certaine valeur propre  $\lambda$  de  $\hat{\mathbf{R}}$ . Et (9) impose que  $\mathbf{v}_1$  soit unitaire. Donc  $\exists d \in \llbracket 1, p \rrbracket$ ,  $\mathbf{v}_1 = \mathbf{u}_d$ . C'est-à-dire qu'un vecteur propre unitaire de  $\hat{\mathbf{R}}$  maximise l'énergie de projection.

$\hat{\mathbf{R}}$  se diagonalise sur la base  $(\mathbf{u}_i)_{i \in \llbracket 1, p \rrbracket}$  avec les valeurs propres  $(\lambda_i)_{i \in \llbracket 1, p \rrbracket}$ . On a donc :

$$\begin{aligned} \mathbf{E}_1 &= \mathbf{u}_d \hat{\mathbf{R}} \mathbf{u}_d^T \\ &= \mathbf{u}_d [\mathbf{u}_1 | \mathbf{u}_2 | \dots | \mathbf{u}_p] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_p^T \end{bmatrix} \mathbf{u}_d^T \\ &= \lambda_d \end{aligned}$$

Comme on souhaite maximiser  $\mathbf{E}_1$ , on prend  $d = 1$  car  $\lambda_1 = \max_{d \in \llbracket 1, p \rrbracket} (\lambda_d)$ . On en déduit donc que  $\mathbf{v}_1 = \mathbf{u}_1$ .

$$\mathbf{u}_1 \in \operatorname{argmax}_{\substack{\mathbf{v}_1 \in \mathcal{R}^p \\ \|\mathbf{v}_1\|_2=1}} \frac{1}{n} \sum_{i=1}^n \langle \mathbf{v}_1, \mathbf{x}_i - \bar{\mathbf{x}} \rangle^2$$

$$\mathbf{E}_1 = \lambda_1$$

## 2.2 Extension du résultat

Montrons pour  $j \in \llbracket 2, p \rrbracket$  que le vecteur propre  $\mathbf{u}_j$  de  $\hat{\mathbf{R}}$  maximise l'énergie de projection dans l'espace orthogonal aux vecteurs  $(\mathbf{u}_i)_{i \in \llbracket 1, j-1 \rrbracket}$ .

Soit  $j \in \llbracket 2, p \rrbracket$ . Avec un raisonnement identique au cas précédent (remplacer les indices 1 par  $j$ ), on arrive au système suivant :

$$\begin{cases} (\hat{\mathbf{R}} - \lambda I_n)\mathbf{v}_j &= 0 \\ \|\mathbf{v}_j\|_2 &= 1 \\ \forall i \in \llbracket 1, j-1 \rrbracket, \langle \mathbf{v}_j, \mathbf{u}_i \rangle &= 0 \end{cases} \quad (10)$$

$$(11)$$

$$(12)$$

(10) et (11) imposent  $\mathbf{v}_j \in \{\mathbf{u}_i | i \in \llbracket 1, p \rrbracket\}$ .

Mais (12) restreint cet ensemble de solution, et  $\mathbf{v}_j \in \{\mathbf{u}_i | i \in \llbracket j, p \rrbracket\}$  car un vecteur ne peut être à la fois orthogonal à lui-même et de norme euclidienne non nulle.

On a alors :

$$\exists d \in \llbracket j, p \rrbracket, \quad \begin{cases} \mathbf{v}_j = & \mathbf{u}_d \\ \mathbf{E}_j = & \lambda_d \end{cases}$$

Comme on souhaite maximiser  $\mathbf{E}_j$  et que  $\lambda_j = \max_{d \in \llbracket j, p \rrbracket} (\lambda_d)$ , on obtient  $d = j$ .

$$\mathbf{u}_j \in \underset{\substack{\mathbf{v}_j \in \mathcal{R}^p \\ \|\mathbf{v}_j\|_2=1 \\ \mathbf{v}_j \perp \mathbf{u}_1, \dots, \mathbf{v}_j \perp \mathbf{u}_{j-1}}}{\operatorname{argmax}} \frac{1}{n} \sum_{i=1}^n \langle \mathbf{v}_j, \mathbf{x}_i - \bar{\mathbf{x}} \rangle^2$$

$$\mathbf{E}_j = \lambda_j$$

## 2.3 Réduction de dimension

Explicitons le ratio de l'énergie de projection  $\mathcal{K}(l)$  en fonction des valeurs propres de  $\hat{\mathbf{R}}$ .

Notons  $\forall i \in \llbracket 1, n \rrbracket, \mathbf{y}_i = \mathbf{x}_i - \bar{\mathbf{x}}$ . On a :

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \|\pi_{\mathcal{S}}(\mathbf{x}_i - \bar{\mathbf{x}})\|_2^2 &= \frac{1}{n} \sum_{i=1}^n \|\pi_{\mathcal{S}}(\mathbf{y}_i)\|_2^2 \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^l \langle \mathbf{u}_j, \mathbf{y}_i \rangle^2 \\ &= \sum_{j=1}^l \frac{1}{n} \sum_{i=1}^n \langle \mathbf{u}_j, \mathbf{y}_i \rangle^2 \\ &= \sum_{j=1}^l \mathbf{E}_j \\ &= \sum_{j=1}^l \lambda_j \end{aligned}$$

En remarquant que pour  $\mathcal{S} = \mathcal{R}^p$ , on a  $\mathbf{y}_i = \pi_{\mathcal{S}}(\mathbf{y}_i)$ , on en déduit :

$$\frac{1}{n} \sum_{i=1}^n \|\mathbf{x}_i - \bar{\mathbf{x}}\|_2^2 = \sum_{j=1}^p \lambda_j$$

Finalement :

$$\mathcal{K}(l) = \frac{\sum_{j=1}^l \lambda_j}{\sum_{j=1}^p \lambda_j}$$

## 2.4 Implémentation

### 2.4.1 Rang de la matrice des données $\mathbf{X}$

Montrons que  $\mathbf{X}$  est de rang  $n-1$ .

$$\begin{aligned} \operatorname{rg}(\mathbf{X}) &= \operatorname{rg}([\mathbf{x}_1 - \bar{\mathbf{x}}, \mathbf{x}_2 - \bar{\mathbf{x}}, \dots, \mathbf{x}_n - \bar{\mathbf{x}}]) \\ &= \operatorname{rg}\left(\left[\frac{1}{n}(\mathbf{x}_1 - \bar{\mathbf{x}}) + \frac{1}{n} \sum_{i=2}^n (\mathbf{x}_i - \bar{\mathbf{x}}), \mathbf{x}_2 - \bar{\mathbf{x}}, \dots, \mathbf{x}_n - \bar{\mathbf{x}}\right]\right) \\ &= \operatorname{rg}\left(\left[\frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}}), \mathbf{x}_2 - \bar{\mathbf{x}}, \dots, \mathbf{x}_n - \bar{\mathbf{x}}\right]\right) \\ &= \operatorname{rg}([\mathbf{0}, \mathbf{x}_2 - \bar{\mathbf{x}}, \dots, \mathbf{x}_n - \bar{\mathbf{x}}]) \end{aligned}$$

A cause de la colonne nulle :  $rg(\mathbf{X}) \leq n - 1$ .

A présent, essayons d'annuler une colonne  $j$  pour  $j \in \llbracket 2, n \rrbracket$ . On cherche donc une suite de scalaire  $(\mu_i)_{i \in \llbracket 2, n \rrbracket}$  tel que :

$$\sum_{i=2}^n \mu_i (\mathbf{x}_i - \bar{\mathbf{x}}) = 0 \quad (13)$$

On dispose de  $(n-1)$  scalaires pour créer une combinaison linéaire nulle des vecteurs  $\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n$  et  $\bar{\mathbf{x}}$ . Or les vecteurs  $\mathbf{x}_2, \dots, \mathbf{x}_n$  sont par hypothèse linéairement indépendant. Et comme  $\bar{\mathbf{x}}$  dépend de  $\mathbf{x}_1$ ,  $\bar{\mathbf{x}}$  est également linéairement indépendant des vecteurs  $\mathbf{x}_2, \dots, \mathbf{x}_n$ . On en conclut que (13) ne peut avoir de solutions.

Il est donc impossible d'annuler une seconde colonne  $j$ . On a donc :  $rg(\mathbf{X}) > n - 2$ .

$$rg(\mathbf{X}) = n - 1$$

On en déduit au passage que  $\lambda_n = 0$  car :

1.  $\hat{\mathbf{R}} = \mathbf{X}\mathbf{X}^T$  n'est pas de rang plein.
2. une matrice de rang non plein admet une valeur propre nulle.
3.  $\lambda_n$  est la plus petite des valeurs propres de  $\hat{\mathbf{R}}$ .
4.  $\hat{\mathbf{R}}$  est semi-définie positive

## 2.5 Lien avec les vecteurs propres de la matrice de Gram

Montrons que si  $\mathbf{v}$  est un vecteur propre de la matrice de Gram  $\mathbf{X}^T\mathbf{X}$ , alors  $\mathbf{u} = \mathbf{X}\mathbf{v}$  est un vecteur propre de  $\hat{\mathbf{R}} = \mathbf{X}\mathbf{X}^T$  associé à la même valeur propre.

Soient  $\lambda$  et  $\mathbf{v}$ , respectivement une valeur propre de  $\mathbf{X}^T\mathbf{X}$  et un vecteur propre associé. On a :

$$(\mathbf{X}^T\mathbf{X} - \lambda I_n)\mathbf{v} = 0$$

Étudions alors le vecteur  $\mathbf{u} = \mathbf{X}\mathbf{v}$  :

$$\begin{aligned} (\mathbf{X}\mathbf{X}^T - \lambda I_n)\mathbf{u} &= (\mathbf{X}\mathbf{X}^T - \lambda I_n)\mathbf{X}\mathbf{v} \\ &= (\mathbf{X}\mathbf{X}^T\mathbf{X} - \lambda\mathbf{X})\mathbf{v} \\ &= \mathbf{X}(\mathbf{X}^T\mathbf{X} - \lambda I_n)\mathbf{v} \\ &= 0 \end{aligned}$$

$\mathbf{u} = \mathbf{X}\mathbf{v}$  est un vecteur propre de  $\hat{\mathbf{R}}$ . Il est associé à la même valeur propre que le vecteur propre  $\mathbf{v}$  de la matrice de Gram  $\mathbf{X}^T\mathbf{X}$ .

## 2.6 Matrice des vecteurs propres orthonormés

Soient  $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n-1}]$  une matrice de vecteurs propres orthonormés de la matrice de Gram  $\mathbf{X}^T\mathbf{X}$  associés à des valeurs propres non nulles. Montrons que  $\mathbf{U} = \mathbf{X}\mathbf{V}(\mathbf{V}^T\mathbf{X}^T\mathbf{X}\mathbf{V})^{-1/2}$  est une matrice à colonnes orthonormées de vecteurs propres de  $\hat{\mathbf{R}}$  associés aux valeurs propres non nulles.

$$\begin{aligned} \mathbf{X}\mathbf{V} &= [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n-1}] \\ \mathbf{V}^T\mathbf{X}^T\mathbf{X}\mathbf{V} &= (\mathbf{X}\mathbf{V})^T(\mathbf{X}\mathbf{V}) = \text{Gram}(\mathbf{X}\mathbf{V}) = \text{Gram}([\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n-1}]) \end{aligned}$$

$$\mathbf{U} = \frac{[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n-1}]}{\sqrt{\text{Gram}([\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n-1}] )}}$$

### 3 Réduction de dimension et eigenfaces

#### 3.1 Les (n-1) eigenfaces

Sur la figure 1, on arrive à distinguer les principales composantes qui vont nous aider à reconstruire les visages, comme les yeux, le nez ou encore la bouche. Le reste du visage semble être quasi commun à tous, puisqu'il ne fait l'objet d'aucune sur-brillance.

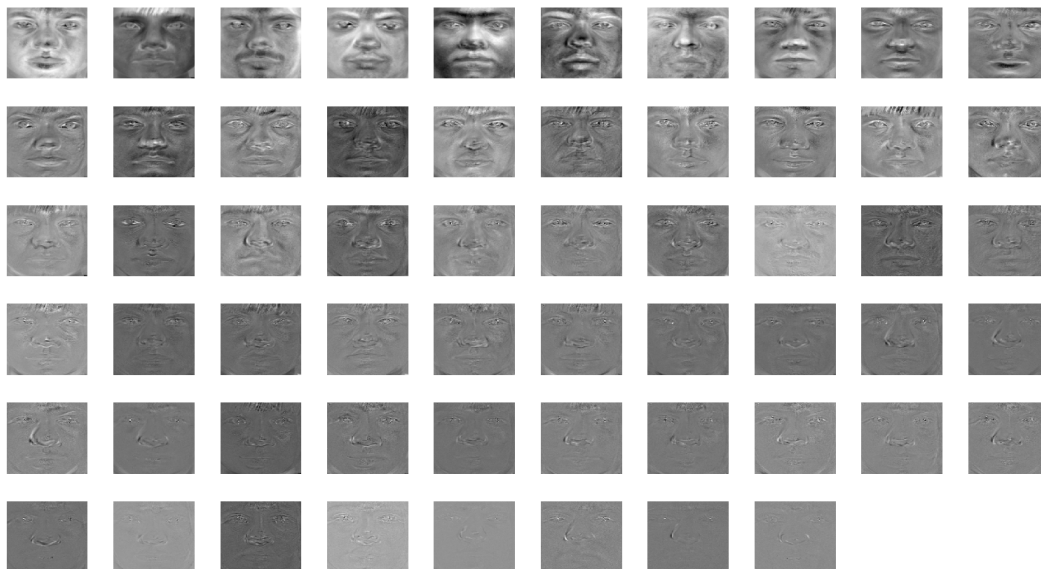


Figure 1: Eigenfaces

En haut à gauche se trouvent les eigenfaces des dimensions les plus énergétiques, et en bas à droite celles ayant le moins d'énergie. On constate une diminution de l'information (visage moins marqué, se confond avec le fond) pour les dimensions les plus faibles.

#### 3.2 Reconstruction d'image après ACP

Plus le nombre de composantes utilisées est élevé, plus l'image après reconstruction par ACP est fidèle à celle d'origine. Ceci paraît cohérent car plus on utilise de composantes, plus on a de possibilités différentes de visage obtenu donc le visage reconstruit est plus susceptible d'être le visage recherché.



Figure 2: Reconstruction par ACP

La figure 2 montre également qu'à partir d'une certaine dimension, le visage est parfaitement reconstruit pour notre oeil. Ici, on ne fait plus la différence entre l'original et une image reconstruite avec les 8 premières dimensions les plus énergétiques.

### 3.3 Ratio de reconstruction

À partir de  $l_{opt}$  composantes, les différences de reconstruction sont minimales donc rien ne sert d'utiliser plus de composantes. En effet, à partir d'un tel ratio de reconstruction, on est censé pouvoir reconnaître le sujet plutôt aisément.

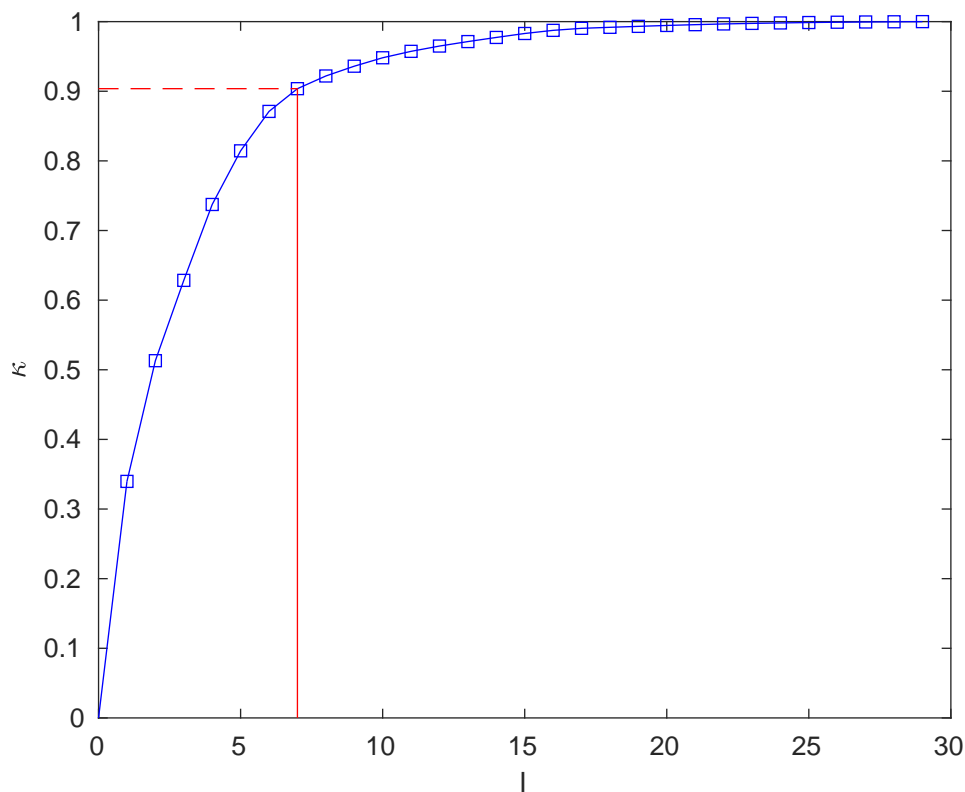


Figure 3: Ratio de reconstruction

En choisissant un ratio de reconstruction de 90 %, on lit sur la figure 3 que les 8 premières dimensions sont suffisantes. La reconstruction avec ces dimensions a été présentée à la section précédente. Et visuellement, on ne fait pas la différence avec l'originale.

### 3.4 Limitation de la seconde base d'entraînement

La base training2 contient 5 images par individu, c'est-à-dire deux fois moins que la base training1. Les deux bases utilisant un éclairage de face, seul le nombre d'images par individu les différencie. Par conséquent, si l'on utilise la base training2, on peut s'attendre à une classification moins pertinente puisque l'étape d'apprentissage sera deux fois plus courte. Le système aura donc deux fois moins d'informations pour classer les futures images des bases test, d'où la classification moins pertinente.

## 4 Classification

### 4.1 Classifieur k-NN

#### 4.1.1 Compromis autour de l'hyperparamètre $k$

Plus le nombre de plus proches voisins est élevé, plus il y a de chance que certains d'entre eux se trouvent dans la mauvaise classe. Par exemple, pour une base de données constituée de 10 photos pour 6 individus (cf. training1), rien ne sert de prendre 15 voisins ( $k = 15$ ) pour effectuer la classification car on sait d'ores et déjà que 5 d'entre eux seront de la mauvaise classe. Par ailleurs, en n'en prenant pas assez, on risque que le peu de voisins trouvés soit faux.

#### 4.1.2 Matrices de confusions

Pour les deux premières bases test, la matrice de confusion est égale à la matrice identité, ce qui est caractéristique d'une classification exacte à 100 %. À partir de la troisième base test, plus on avance dans les bases, plus l'angle d'éclairage est fort donc moins les caractéristiques globales de chacun des individus sont reconnaissables. La classification devient nettement plus complexe. Cela se traduit par la dégradation de la matrice de confusion et du taux d'erreur global.

### 4.2 Classifieur Gaussien

#### 4.2.1 Choix du modèle gaussien

Sur la figure 4, les photos des trois individus sont représentées par des \* de même couleur, formant des clusters. Les moyennes intra-classes associées sont représentées par des points plus grand et de même couleur.

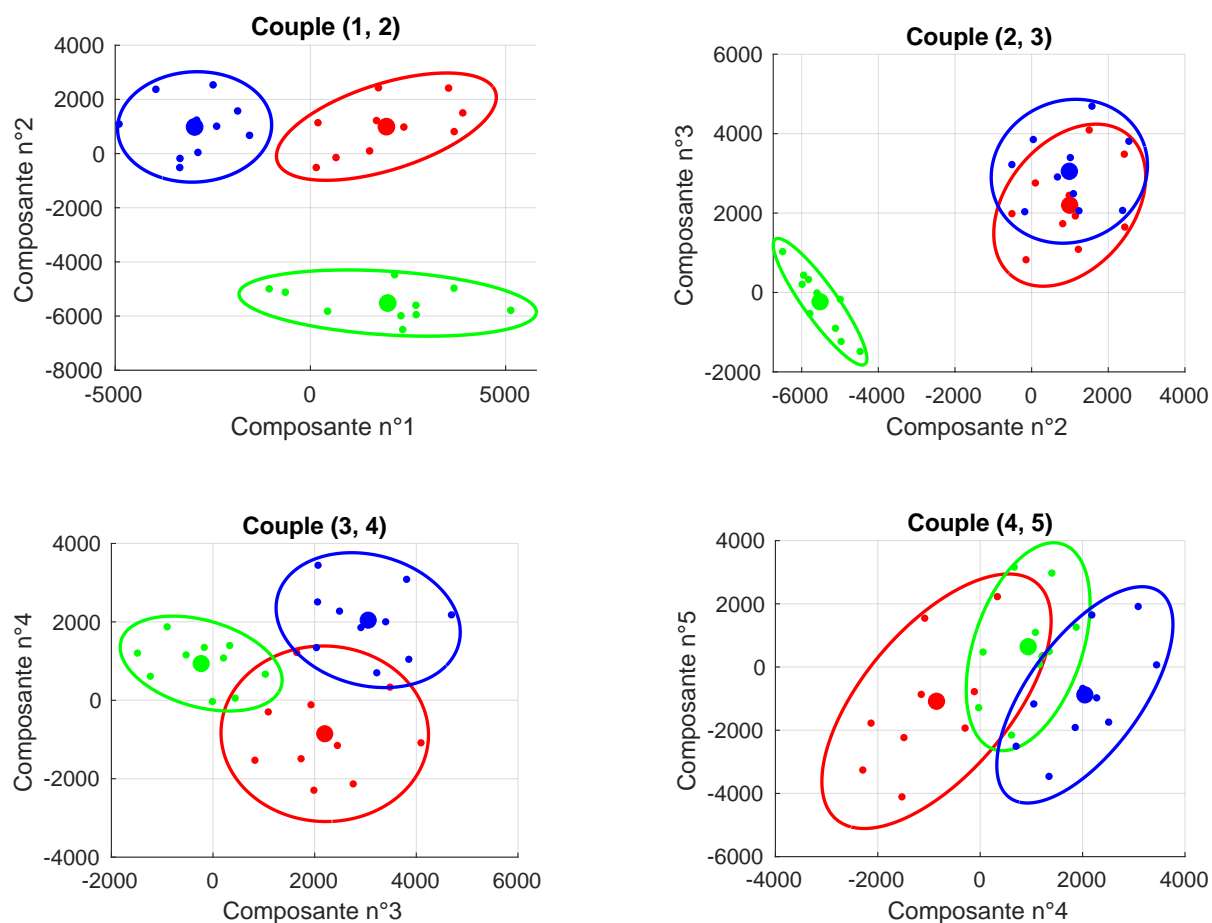


Figure 4: Les premières composantes principales pour les trois premiers individus

Les clusters semblent se répartir de manière plutôt circulaire autour de la moyenne, ce qui justifie le choix d'un modèle gaussien.

#### 4.2.2 Limitation potentielle du classifieur gaussien

Dans le modèle gaussien, nous sommes amenés à inverser la matrice  $\Sigma$ . Pour éviter des phénomènes de divergence, cette matrice se doit d'être de rang plein. Seulement :

$$\begin{cases} \dim(\Sigma) = (l, l) \\ \text{rg}(\Sigma) \leq \min(l, n - m) \end{cases}$$

On en déduit qu'il est nécessaire que

$$l \leq n - m$$

Autrement dit, le nombre de dimensions que l'on peut étudier ( $l$ ) n'est pas limité par le nombre d'images possédées ( $n$ ) mais par le nombre d'image possédées et le nombre de classes ( $n-m$ ). Ce qui nous ramène au nombre de photos par individus. Il doit être suffisamment important pour que ce classifieur puisse fonctionner.

### 4.3 Comparaison des classifieurs k-NN et Gaussien

Les performances présentées en figure 5 se dégradent au fur et à mesure car les bases de tests sont de moins en moins bonnes. Mais on s'aperçoit que le classifieur gaussien est plus robuste que le classifieur k-NN.

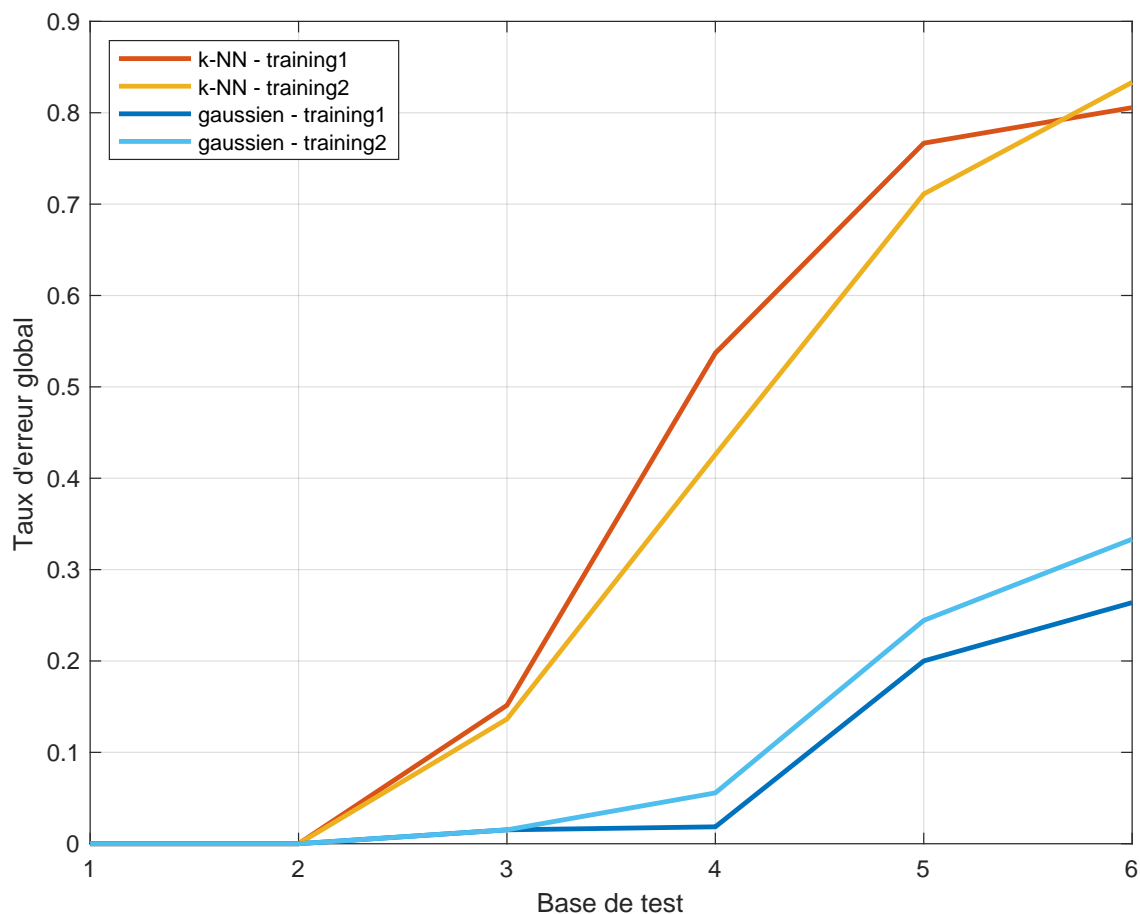


Figure 5: Performance des deux classifieurs



#### 4.4 Amélioration du classifieur Gaussien

Le but est de détecter des visages n'appartenant pas à la base d'entraînement. Pour cela, on étudie les ellipses présentées en figure 4 qui forment des clusters autour des visages d'entraînement associés.

Lorsqu'un visage est testé, on regarde s'il est à l'intérieur d'une ellipse. Si oui, son score au visage associé à l'ellipse est incrémenté. Une fois que toutes les dimensions de plus fortes énergies ont été étudiées, on attribue au visage testé l'identité du visage d'entraînement pour lequel le score a été le plus élevé.

De cette manière, on associe systématiquement un visage à un de la base d'entraînement. Pour y remédier, on impose un score minimal à obtenir au visage test. S'il n'est pas atteint alors on décide que le visage étudié n'appartient pas à la base d'entraînement.

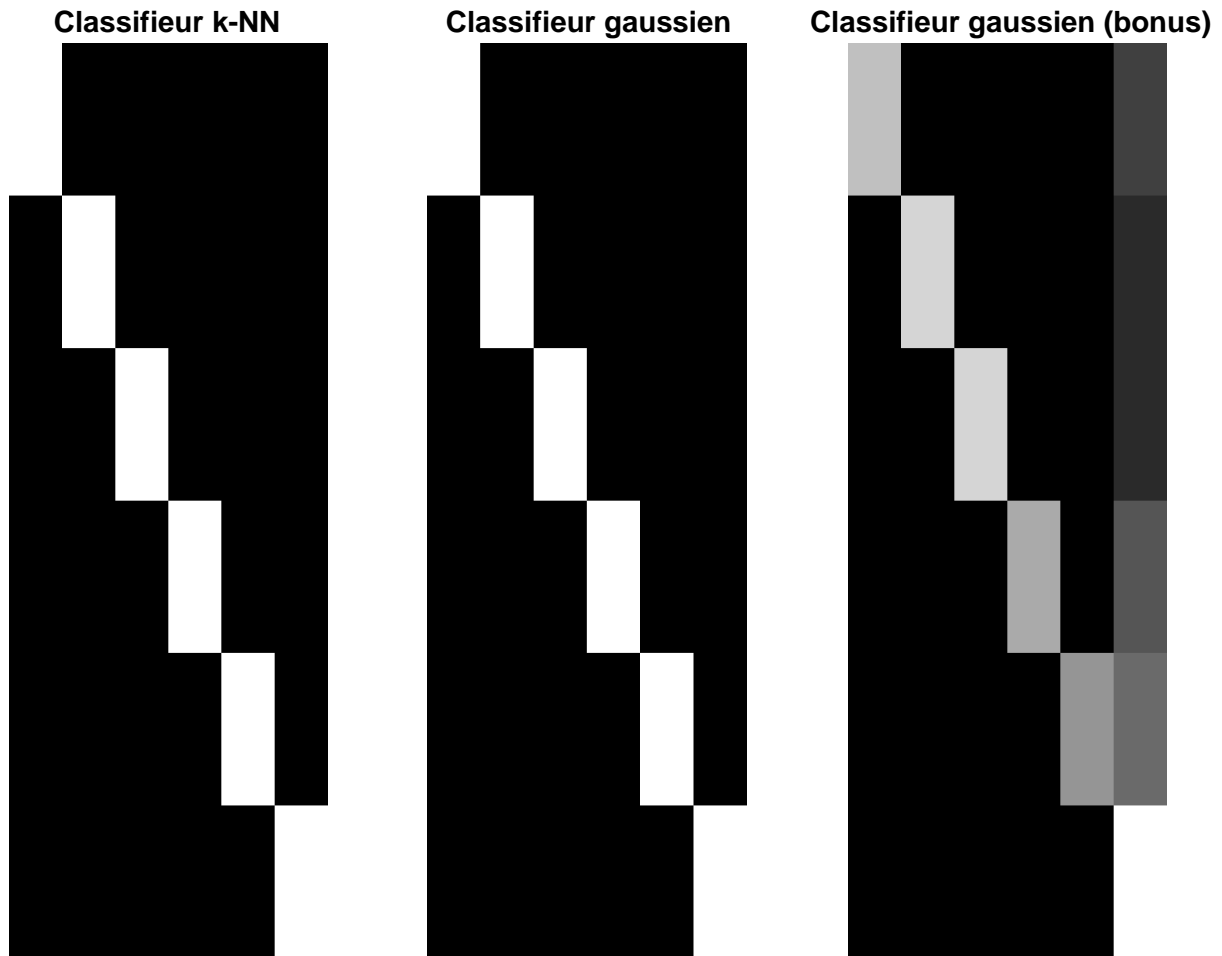


Figure 6: Matrices de confusion

Les premières composantes sont celles qui concentrent le plus d'énergie, donc celles qui permettent le plus de caractériser un visage en général. C'est pourquoi plus on avance dans les composantes, plus les clusters se superposent puisque les composantes deviennent de moins en moins caractéristiques et de plus en plus communes à tous les visages.