

Projet 3A
Analyse d'imagerie polarimétrique

GUINAUDEAU Alexandre HULOT Pierre DEJOIE Etienne

6 février 2016

Résumé

Le résumé (abstract en anglais) de mon article.

Chapitre 1

contexte

1.1 ADM Polar, contexte du projet

1.2 Presentation des données

A mettre ici : présentation du set données. Nombre de pièces différentes.
taille des images. Explication de ce qu'est la matrice de Muller

Chapitre 2

Le traitement des données

2.1 Prétraitement des données

Les différents types de prétraitements que l'on peut faire avant de traiter les données

2.2 Les différentes approches de traitement des données

2.2.1 Réduction de dimension

PCA

rappel de la méthode L'Analyse en Composantes Principales (ou PCA) consiste à essayer de représenter les données dans un espace de plus petites dimensions. Les vecteurs directeurs du nouvel espace maximise la variance entre les données. Nous présentons ici les résultats pour la dimension 2.

prétraitement utilisé Nous effectuons cette PCA sur les centres des clusters préalablement présentés (cf 1.1.1). Les centres des clusters représentent de manière fidèle l'ensemble des points qu'il rassemble. Chaque cluster est représenté par un vecteur d'éléments de la matrice de Muller. Tous les éléments de Muller sont gardés à l'exception de la première ligne et première colonne qui ne sont pas a priori pertinentes (d'après les informations des physiciens)

résultats La réduction de dimension par PCA semble efficace. La composante principale explique 90% de la variance (fig 2.3). De l'analyse de la première composante (fig 2.4) ressort deux effets principaux : - La petite contribution des éléments diagonaux de la matrice de Muller - Le rôle prépondérant de M34 et M43

On remarque une certaine anticorrelation des éléments de la matrice de Muller. Le poids de M43 est proche de l'opposé de celui de M34. Le poids de M42 est également proche de l'opposé de celui de M24. Cette observation n'est par contre pas vérifiée pour M23 et M32 qui semblent corrélés.

FIGURE 2.1 – centre des clusters avant transformation

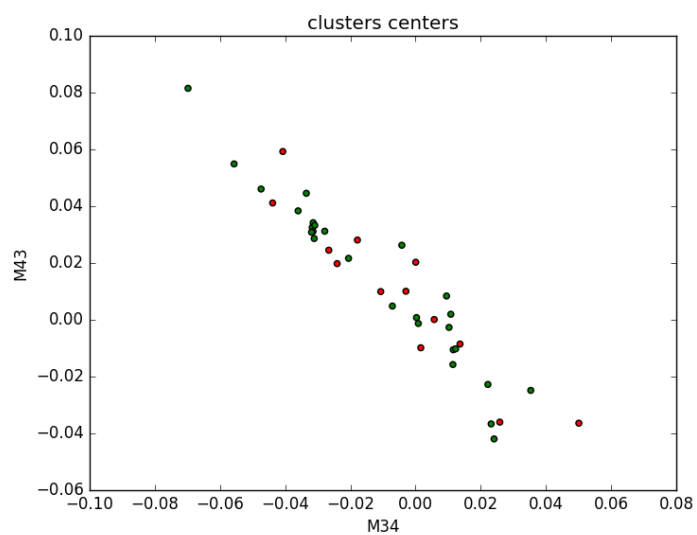


FIGURE 2.2 – centre des clusters après transformation

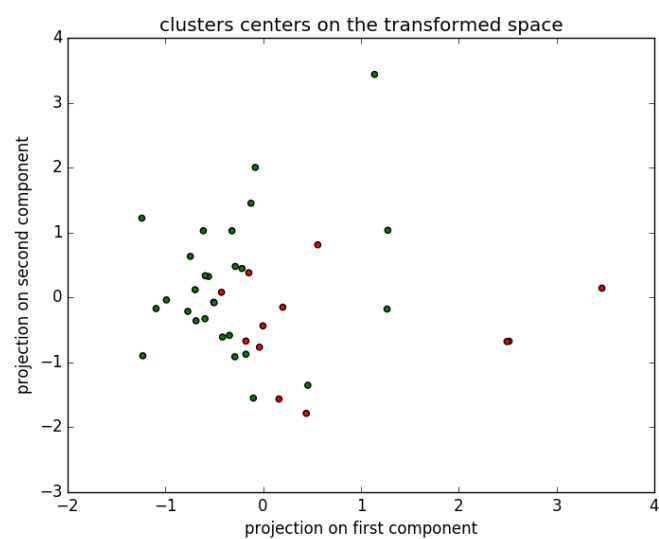


FIGURE 2.3 – Part de variance expliquée

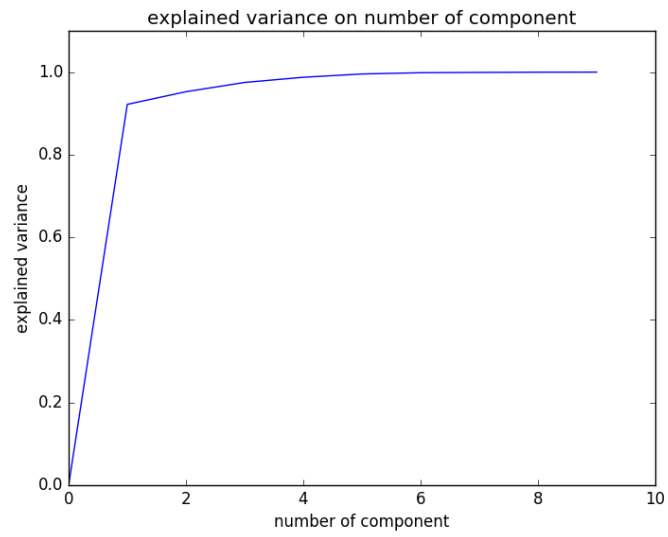
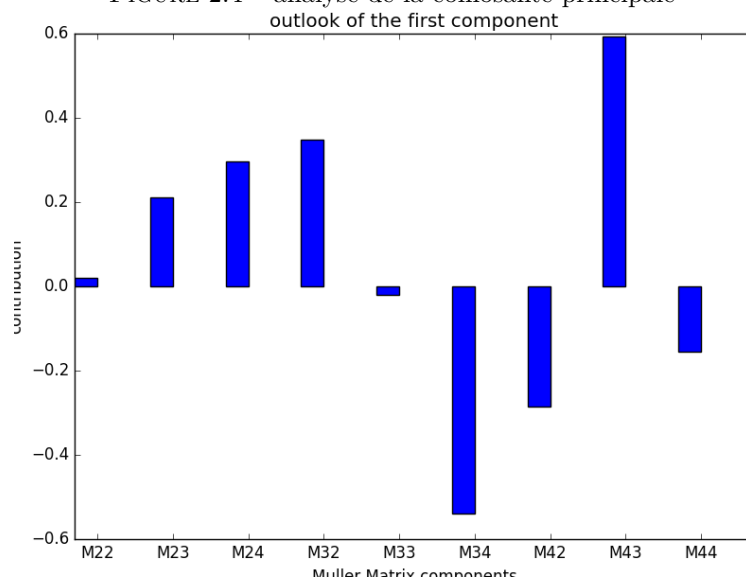


FIGURE 2.4 – analyse de la comosante principale



conclusion La PCA effectuée sur les centres des clusters valident certaines supposition comme le rôle faible des éléments diagonaux ou le rôle important des éléments M34 et M43.

Par contre, la projection de la PCA en 2 dimensions ne nous permet pas de séparer les données de manière suffisante pour être capable de distinguer des zones clairement différentes entre les clusters sains et les clusters malades. (fig 2.2)

2.2.2 Méthode de classification

Arbre décisionnel et Random Forest

rappel de la méthode

prétraitement utilisé

résultats (notamment graphique)

explication

piste d'amélioration

K plus proche voisin

rappel de la méthode La méthode des k plus proche voisin consiste à essayer de prédire l'état d'un nouveau point en se basant sur l'état de ses voisins les plus proches. Cette méthode a l'avantage de pouvoir classer les données selon des schémas non linéaires. Par contre, c'est une méthode sensible à la dimension.

prétraitement utilisé Nous avons testé la méthode des knn sur les centres des clusters pour plusieurs raisons : - A chaque classification, l'algorithme doit recalculer l'ensemble des distances avec tout l'échantillon d'apprentissage. Il y a donc une nécessité de réduire le nombre de données en entrée sur lesquels on calcule les distances. - De plus, chaque cluster regroupe un ensemble de points très proches les uns des autres. Ainsi, en prenant tous les éléments, les plus proches voisins d'un certain point auraient souvent tous appartenu au même cluster ce qui rend l'information extraite redondante. Cela aurait valu à utiliser l'algorithme avec 1 seul voisin.

résultats

piste d'amélioration

Table des matières

1	contexte	1
1.1	ADM Polar, contexte du projet	1
1.2	Presentation des données	1
2	Le traitement des données	2
2.1	Prétraitement des données	2
2.2	Les différentes approches de traitement des données	2
2.2.1	Réduction de dimension	2
2.2.2	Méthode de classification	5
	Bla bla bla	

Table des figures

2.1	centre des clusters avant transformation	3
2.2	centre des clusters après transformation	3
2.3	Part de variance expliquée	4
2.4	analyse de la comosante principale	4

Liste des tableaux