

# **Reconnaissance de chiffres et débruitage par généralisation de l'ACP aux RKHS**

El Bouzekraoui Younes — MDAA Saad

Département Sciences du Numérique - Deuxième année  
2020-2021

# Contents

<b>1</b>	<b>Stratégie 1: ACP classique</b>	<b>3</b>
1.1	Sensibilité des résultats . . . . .	3
<b>2</b>	<b>Stratégie 2: généralisation aux RKHS</b>	<b>4</b>
2.1	Noyau linéaire . . . . .	4
2.2	Noyau Gaussien . . . . .	6
2.3	Sensibilité des résultats . . . . .	6

# 1 Stratégie 1: ACP classique

## 1.1 Sensibilité des résultats

On remarque que pour une variance de bruit importante ( $sig0 = 0.4$ ) si on choisit une précision proche de 1, on obtient une classification correcte mais les résultats de la reconstruction ne sont pas satisfaisantes, en effet avec une précision proche de 1 on ne réduit pas assez la dimension du problème et donc on n'élimine pas des vecteurs propres ( $u_i$ ) qui contiennent majoritairement que du bruit lors de la reconstruction.

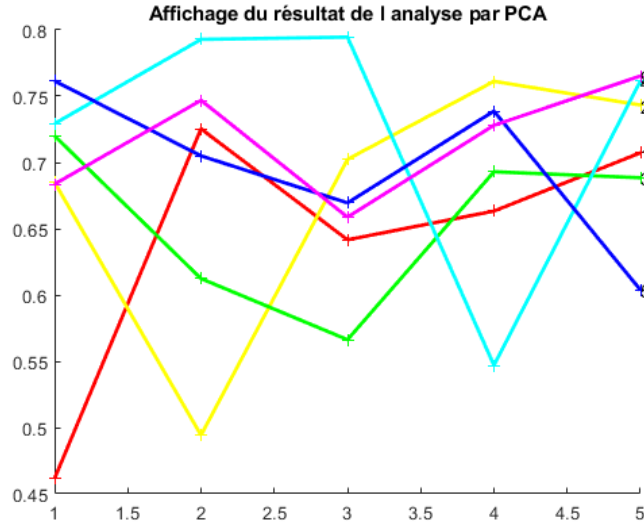


Figure 1: résultats de la classification par PCA pour  $sig0 = 0.4$  et  $Precapprox = 0.9$

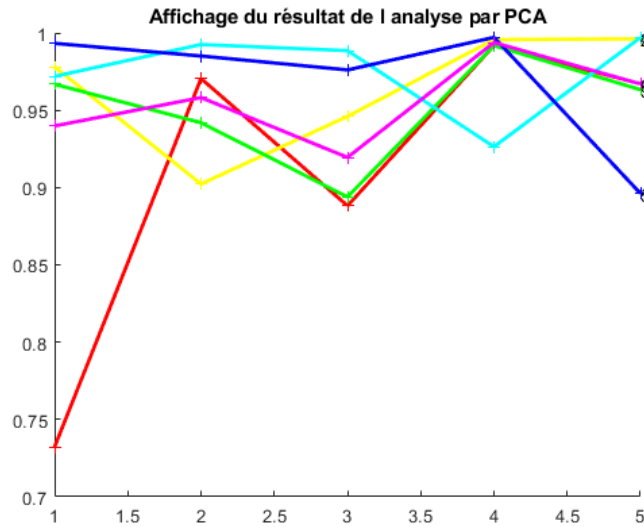


Figure 2: résultats de la classification par PCA pour  $sig0 = 0.4$  et  $Precapprox = 0.2$

- On remarque que le chiffre 9 est proche du chiffre 3 ce qui est logique.

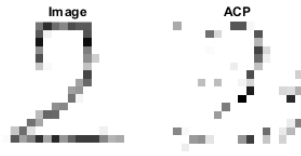


Figure 3: reconstruction par PCA pour  $sig0 = 0.4$  et  $Precapprox = 0.9$

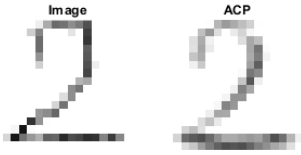


Figure 4: reconstruction par PCA pour  $sig0 = 0.4$  et  $Precapprox = 0.2$

## 2 Stratégie 2: généralisation aux RKHS

### 2.1 Noyau linéaire

On constate que l'algorithme kacp est beaucoup plus rapide que l'algorithme acp (0.5 sec vs 0.06 sec)

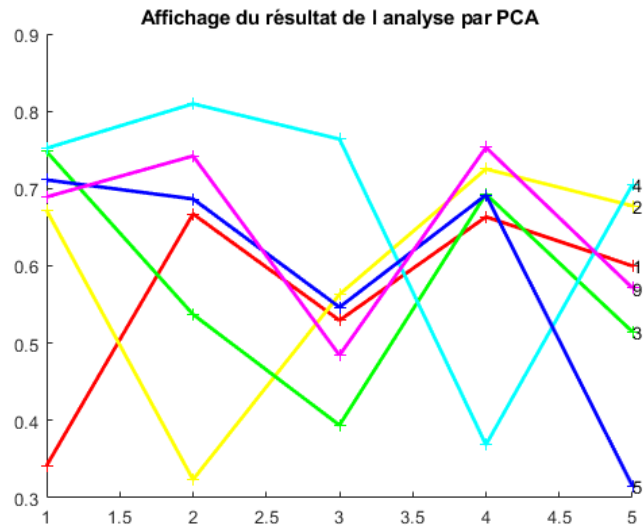


Figure 5: résultats de la classification par PCA pour  $sig0 = 0.05$  et  $Precapprox = 0.9$

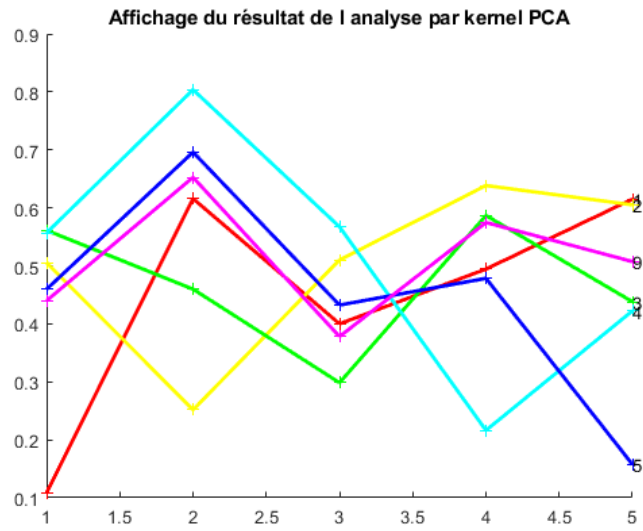


Figure 6: résultats de la classification par KPCA pour  $sig0 = 0.05$  et  $Precapprox = 0.9$

## 2.2 Noyau Gaussien

## 2.3 Sensibilité des résultats

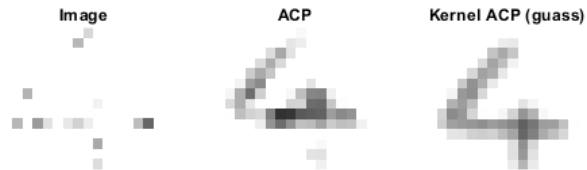


Figure 7: reconstruction par PCA et KPCA pour  $\sigma_0 = 0.4$  et  $Precapprox = 0.2$

- On remarque que l'Acp au noyau Gaussien est très sensible au bruit plus, il est élevé plus la classification tend à être erroné.
- Et concernant la précision, en l'augmentant, on aura plus de composantes principales et donc on aura une bonne classification.
- pour le noyau gaussien on obtient une bonne classification pour  $\sigma = 5$  si on continue à augmenter  $\sigma$  les données se déforment et on obtient une classification erronée (idem pour le noyau polynomial).