



ÉCOLE
CENTRALE LYON

Apprentissage bayésien & Exploration de textes

Rapport de BE

Apprentissage bayésien et chaîne de Markov cachée

MOS 3.1

Étudiants :

MORCHOISNE Dewone

MAÏCHE Ines

Enseignant :

DERRODE Stéphane

15 MARS 2022

Sommaire

1	Introduction	2
2	Restauration non supervisée	2
3	Application à la segmentation d'une image	4

1 Introduction

L'objectif de ce second BE est de mettre en place les techniques d'apprentissage bayésien sur des chaînes de Markov cachées.

2 Restauration non supervisée

A partir uniquement des observations Y d'une chaîne de Markov cachée et du nombre de classes K , il est possible de reconstruire le modèle de Markov associé à X et de retrouver les paramètres des lois gaussiennes associées aux différents états à l'aide de l'algorithme Expectation - Maximization. La classification des observations se fait selon une décision bayésienne.

Au bout de 30 itérations l'algorithme EM donne les résultats suivants :

```
Confusion matrix for MPM =
[[372.  7.]
 [ 10. 611.]]
Global Error rate for MPM : 0.017
Class Error rate for MPM : [0.01846966 0.01610306]
```

On voit bien que l'on est dans un modèle à 2 états. La matrice de confusion nous permet d'observer que le nombre de vrais positifs et vrais négatifs sont très largement supérieurs aux nombres de faux positifs et faux négatifs. Les résultats sont donc assez bons, seul 1.7% des points sont mal classifiés, le taux d'erreur est presque identique pour les 2 classes.

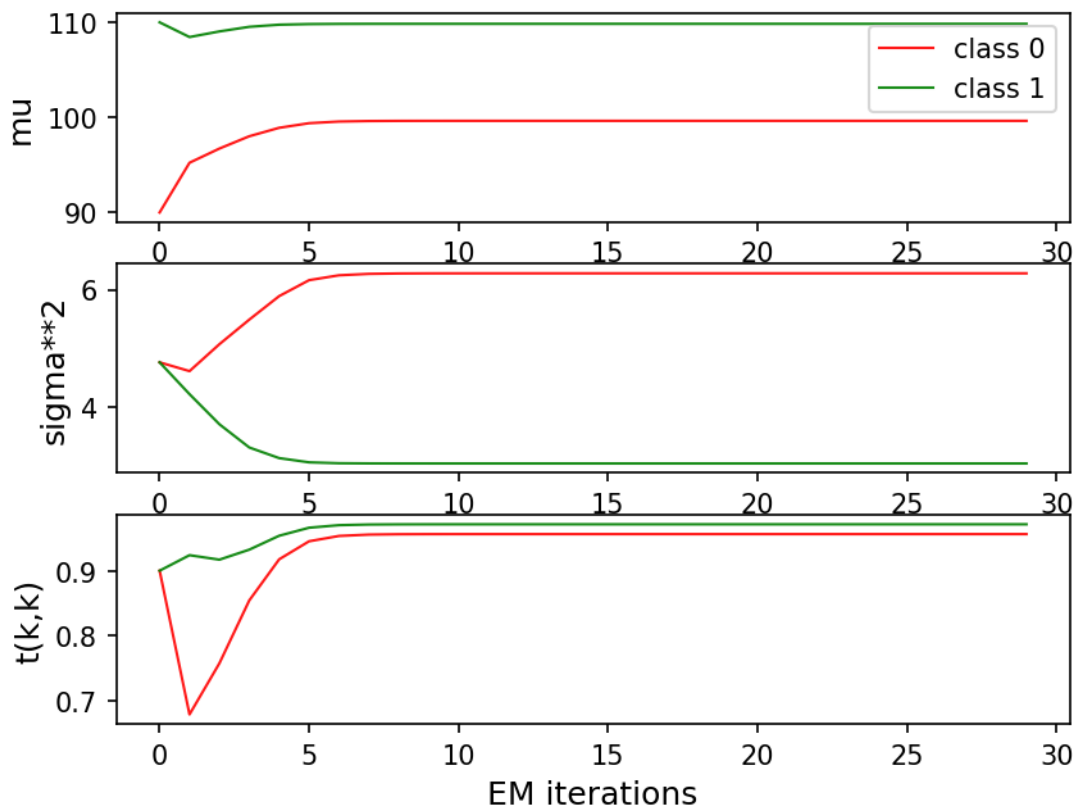


FIGURE 1 – Évolution de l'estimation des paramètres du modèle de Markov X

On peut observer l'évolution de l'estimation des paramètres du modèle au cours de l'exécution de l'algorithme EM. Les paramètres convergent assez rapidement et on retrouve bien les paramètres de 2 gaussiennes de la chaîne de Markov cachée qui avait été simulée lors d'une question précédente. Les légères différences s'expliquent dans un premier temps par le fait que ces paramètres sont obtenus à partir d'un nombre fini d'observations créées par une simulation de la chaîne de Markov cachée originelle.

Paramètre	Valeur Estimée	Valeur avant simulation
Mean	[99.6, 109.8]	[100, 110]
Std	[6.2, 3.1]	[6, 3]
t	[0.956, 0.0435] [0.0282, 0.971]	[0.95, 0.05] [0.05 0.95]

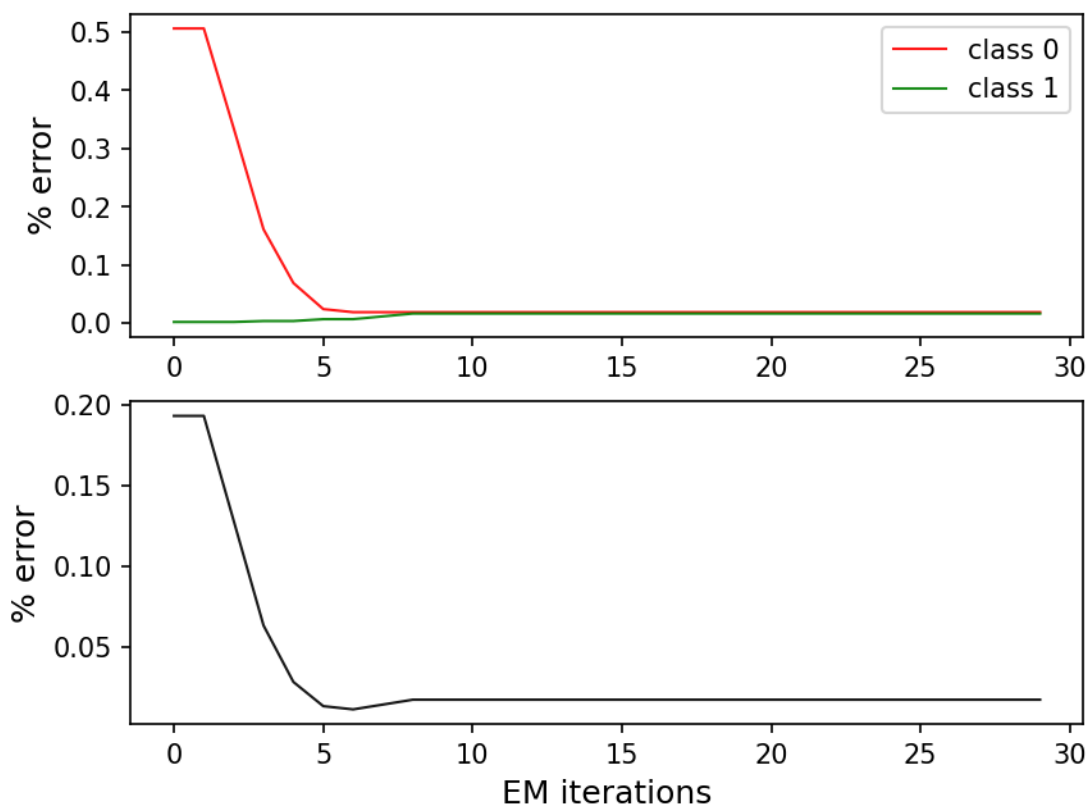


FIGURE 2 – Évolution du taux d'erreur de classification

Les courbes de taux d'erreur mettent en évidence le fait que dans un tout premier temps, tous les points sont classifiés comme faisant partie de la classe 0. On a 50% d'erreur de classification à l'itération 0 et 0% pour la classe 1. Au fur et à mesure de l'apprentissage des paramètres, de plus en plus de points sont classifiés comme appartenant à la classe 1 ce qui réduit le taux d'erreur de la classe 0 et augmente celui de la classe 1. Le taux d'erreur global continue lui de diminuer. Un point intéressant étant le fait que le taux d'erreur global minimum est atteint à la 5ème itération mais qu'il ré-augmente à partir de la 6ème avant de converger.

3 Application à la segmentation d'une image

La méthode de reconstruction d'une chaîne de Markov cachée peut être appliquée pour réaliser de la segmentation d'image. En effet, seuls un vecteur d'observation Y et le nombre de classes souhaitées sont nécessaires. Pour ce faire, on transforme une image 2D en niveau de gris en un vecteur 1D en utilisant le parcours de Peano. On applique ensuite l'algorithme de restauration non supervisé vu précédemment à ce vecteur ce qui va permettre de classifier l'ensemble des observations. On réalise ensuite le parcours inverse de Peano pour reconstruire l'image puis on réajuste les valeurs pour qu'elles soient en niveau de gris 255.

Augmenter le nombre de classes permet d'avoir plus de détails dans l'image mais le temps de calcul augmente asymptotiquement de manière quadratique avec le nombre de classes.



FIGURE 3 – Image originale 64x64 en 255 niveau de gris



FIGURE 4 – Image après classification $K=2$



FIGURE 5 – Image après classification $K=4$



FIGURE 6 – Image après classification $K=8$

On observe au niveau des courbes d'évolution de l'estimation des paramètres du modèle que la convergence se fait en général assez rapidement pour μ , un peu plus lentement pour σ et enfin les $t(k,k)$ sont encore plus lent à converger. Dans le cas de 2 classes les courbes de μ et σ ont convergé ce qui n'est pas le cas des courbes $t(k,k)$. Il faudrait augmenter encore un peu le nombre d'itérations pour mieux reconstituer le modèle. Dans le cas $K=4$, seul les courbes de μ ont convergé. On observe également que les μ des différentes classes sont espacés de manière assez régulière.

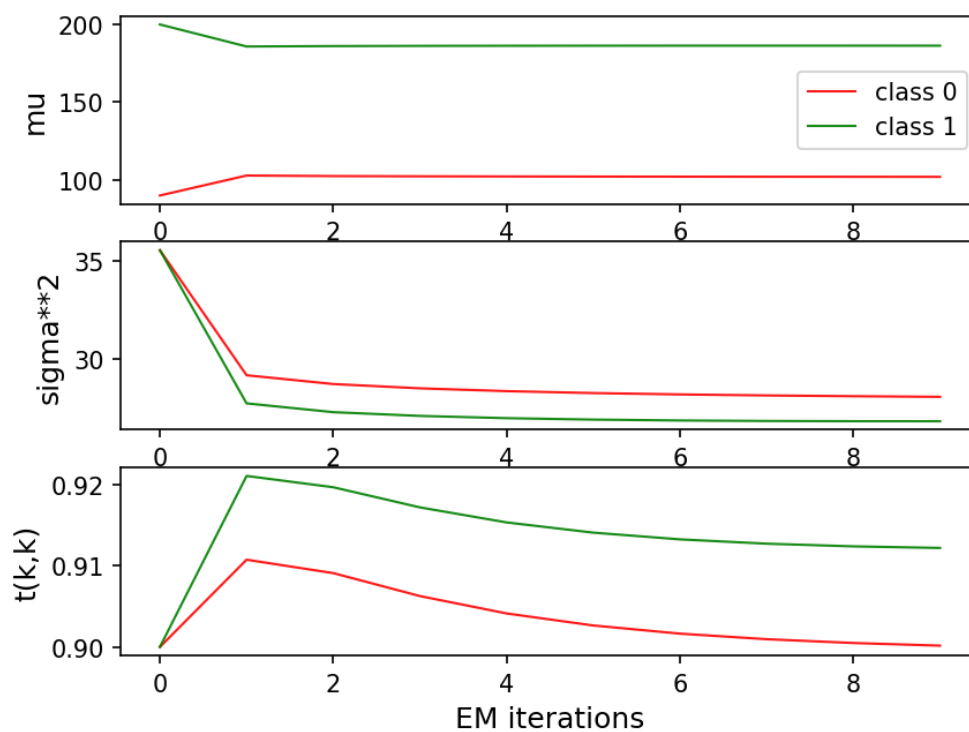


FIGURE 7 – Évolution de l'estimation des paramètres du modèle de Markov X pour $K=2$

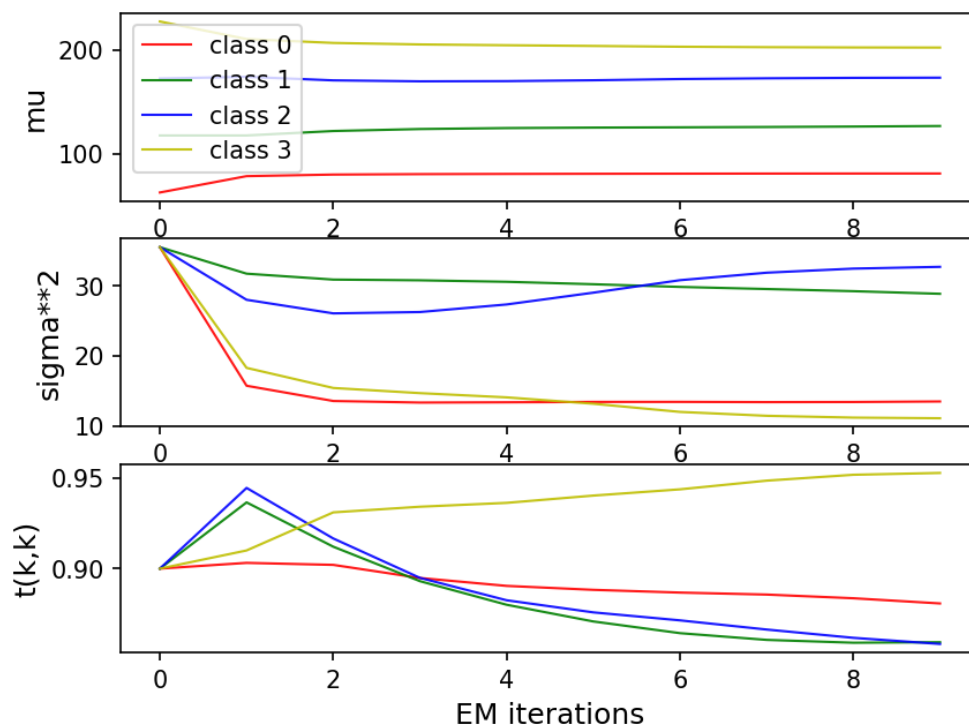


FIGURE 8 – Évolution de l'estimation des paramètres du modèle de Markov X pour $K=4$