

Projet de classification non supervisée

à rendre au plus tard le 01/12/2023

1 Objectifs

L'objectif de ce projet est d'utiliser la classification non supervisée pour résumer les variations de la consommation d'énergie de 100 appartements, observée toutes les 30 minutes durant 91 jours consécutifs. Plus spécifiquement, on cherche ici à obtenir une classification des jours, uniforme pour l'ensemble des appartements. Une telle classification pourra être utile, par exemple, dans le cadre de la surveillance d'un réseau d'électricité.

2 Données

Pour faciliter leur manipulation, les données sont fournies via les trois tableaux suivants, constitués du même nombre de lignes :

- **X** (100×91 lignes et 48 colonnes) : chaque ligne correspond à la consommation observée sur une journée avec un pas de temps de 30 minutes :

```
X = as.matrix(read.table("http://allousame.free.fr/mls/donnees/X.txt",header=F))
```

- **APPART** (100×91 lignes et 1 colonne) : chaque cellule correspond au numéro de l'appartement :

```
APPART = as.matrix(read.table("http://allousame.free.fr/mls/donnees/APPART.txt",header=F))
```

- **JOUR** (100×91 lignes et 1 colonne) : chaque cellule correspond au numéro de la journée :

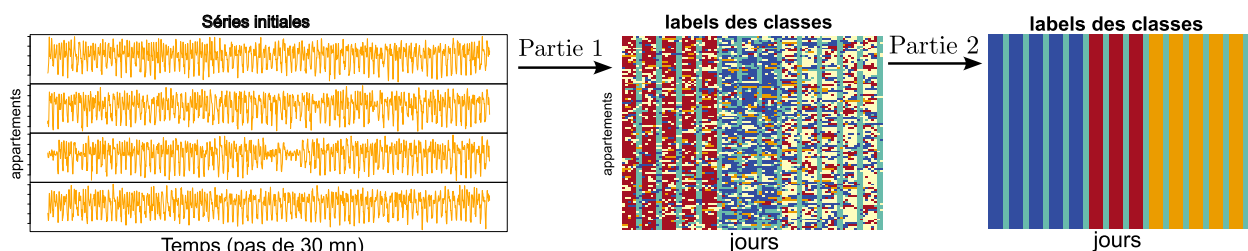
```
JOUR = as.matrix(read.table("http://allousame.free.fr/mls/donnees/JOUR.txt",header=F))
```

Par exemple, pour accéder aux données de l'appartement 1 pour la journée 2, on pourra utiliser la commande

```
plot(X[APPART==1 & JOUR==2,],type="l",lty=1)
```

Pour faciliter l'analyse, chacune des séries a été passée à l'échelle logarithmique, et centrée.

3 Travail à effectuer



Partie 1 : Transformation du tableau X en un tableau de données catégorielles Y

L'objectif de cette partie est de regrouper les lignes du tableau X en K classes, puis, à partir des labels des classes, de transformer les données journalières initiales (X) en un tableau de données catégorielles Y (100 lignes et 91 colonnes). Réaliser cette première étape (voir figure 1) en vous basant sur l'une des méthodes étudiées en cours. Interpréter les classes obtenues ainsi que leurs centres associés.

Partie 2 Réaliser maintenant une classification des 91 jours en partant du tableau de données catégorielles Y . Pour réaliser cette classification, vous pourrez utiliser directement les données catégorielles ou les convertir d'abord en données quantitatives. Analyser/Commenter les résultats obtenus.

4 Documents à rendre

Deux fichiers : un rapport en pdf n'excédant pas 10 pages et un fichier contenant le code (R ou Python) associé.