Assignment 3 (ML for TS) - MVA 2022/2023

Thibault de Surrel thibault.de-surrel@ensta-paris.fr
Alexandre Pradeilles alexandre.pradeilles@gmail.com

April 13, 2023

1 Introduction

Objective. The goal is to implement (i) a signal processing pipeline with a change-point detection method and (ii) wavelets for graph signals.

Warning and advice.

- Use code from the tutorials as well as from other sources. Do not code yourself well-known procedures (e.g. cross validation or k-means), use an existing implementation.
- The associated notebook contains some hints and several helper functions.
- Be concise. Answers are not expected to be longer than a few sentences (omitting calculations).

Instructions.

- Fill in your names and emails at the top of the document.
- Hand in your report (one per pair of students) by Friday 7th April 11:59 PM.
- Rename your report and notebook as follows:
  `FirstnameLastname1_FirstnameLastname1.pdf` and
  `FirstnameLastname2_FirstnameLastname2.ipynb`.
  For instance, `LaurentOudre_CharlesTruong.pdf`.
- Upload your report (PDF file) and notebook (IPYNB file) using this link: https://www.dropbox.com/request/rmETjrLAH9Li3pf8JvOt.

2 Dual-tone multi-frequency signaling (DTMF)

In the last tutorial, we started designing an algorithm to infer from a sound signal the sequence of symbols encoded with DTMF.

Question 1

Finalize this procedure–in particular, find the best hyperparameters. Describe in 5 to 10 lines your methodology and the calibration procedure (give the hyperparameter values).

Answer 1

After the tutorial, we had to calibrate the hyper-parameter "n bkps". This parameter has to be a function of the signal size because depending on the number of symbols we want to detect we cannot define a fixed number of ruptures. After looping from $\frac{signalsize}{20}$ to $\frac{signalsize}{2}$, we kept $nbkps = \frac{signalsize}{6}$ because it was the largest parameter always accepted by the algorithm and it ensure to separate the symbols by empty parts. Then for each part of the signal, we extracted the two main frequencies among the 8 interesting ones and assign the corresponding symbol.

Finally, we kept only the symbols with an average value over 0.25 (parameter chosen based on a loop which measure the precision of the final prediction for different values of this parameter). Then we can regroup the following parts detecting the same symbol because the choice of the rupture parameter ensure an empty part between symbols.

Question 2

What are the two symbolic sequences encoded in the provided signals?

Answer 2

- Sequence 1: B94B38B#1
- Sequence 2: CD112639

3 Wavelet transform for graph signals

Let G be a graph defined a set of n nodes V and a set of edges E. A specific node is denoted by v and a specific edge, by e. The eigenvalues and eigenvectors of the graph Laplacian L are $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and $u_1, u_2, \ldots, u_n$ respectively.

For a signal $f \in \mathbb{R}^n$, the Graph Wavelet Transform (GWT) of f is $W_f : \{1, \ldots, M\} \times V \longrightarrow \mathbb{R}$:

$$W_f(m, v) := \sum_{l=1}^{n} \hat{g}_m(\lambda_l) \hat{f}_l u_l(v) \tag{1}$$

where $\hat{f} = [\hat{f}_1, \ldots, \hat{f}_n]$ is the Fourier transform of f and $\hat{g}_m$ are M kernel functions. The number M of scales is a user-defined parameter and is set to $M := 9$ in the following. Several designs are available for the $\hat{g}_m$; here, we use the Spectrum Adapted Graph Wavelets (SAGW). Formally, each kernel $\hat{g}_m$ is such that

$$\hat{g}_m(\lambda) := \hat{g}^U(\lambda - am) \quad (0 \leq \lambda \leq \lambda_n) \tag{2}$$

where $a := \lambda_n / (M + 1 - R)$,

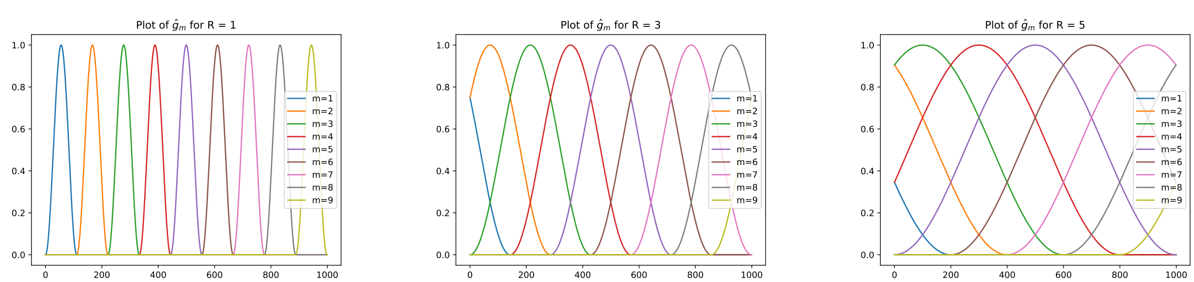
$$\hat{g}^U(\lambda) := \frac{1}{2}\left[1 + \cos\left(2\pi\left(\frac{\lambda}{aR} + \frac{1}{2}\right)\right)\right] \mathbb{1}(-Ra \leq \lambda < 0) \tag{3}$$

and $R > 0$ is defined by the user.

Question 3

Plot the kernel functions $\hat{g}_m$ for $R = 1$, $R = 3$ and $R = 5$ (take $\lambda_n = 12$) on Figure 1. What is the influence of R?

Answer 3



(a) $R = 1$ (b) $R = 3$ (c) $R = 5$

Figure 1: The SAGW kernels functions

We can notice that when R increase, the frequency decrease and the support of the function increase. Therefore, we can notice that the overlap of the filters increases too. For $R = 1$, there is no overlapping as for $R = 5$, there is a lot more overlapping between the different kernel functions $\hat{g}_m$.

We will study the Molene data set (the one we used in the last tutorial). The signal is the temperature.

Question 4

Construct the graph using the distance matrix and exponential smoothing (use the median heuristics for the bandwidth parameter).

- Remove all stations with missing values in the temperature.
- Choose the minimum threshold so that the network is connected and the average degree is at least 3.
- What is the time where the signal is the least smooth?
- What is the time where the signal is the smoothest?

Answer 4

The stations with missing values are :

- ARZAL
- BREST-GUIPAVAS
- BRIGNOGAN
- LANDIVISIAU
- LANNAERO
- LANVEOC
- OUESSANT-STIFF
- PLOUAY-SA
- CAMARET
- BEG-MEIL
- BATZ
- PLOUDALMEZEAU
- PLOUGONVELIN
- QUIMPER
- RIEC SUR BELON
- SIZUN
- ST NAZAIRE-MONTOIR
- VANNES-MEUCON

The maximum threshold so that the network is connected and the average degree is at least 3 is 0.83.

The time where the signal is the smoothest is 2014-01-24 19:00:00. The time where the signal is the least smooth is 2014-01-10 09:00:00.

Question 5

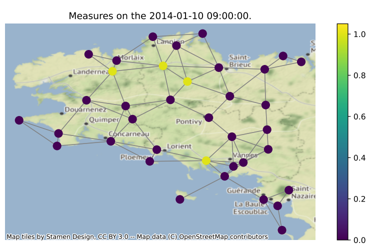
(For the remainder, set $R = 3$ for all wavelet transforms.)

For each node v, the vector $[W_f(1, v), W_f(2, v), \ldots, W_f(M, v)]$ can be used as a vector of features. We can for instance classify nodes into low/medium/high frequency:

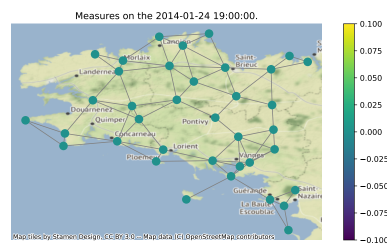
- a node is considered low frequency if the scales $m \in \{1, 2, 3\}$ contain most of the energy,
- a node is considered medium frequency if the scales $m \in \{4, 5, 6\}$ contain most of the energy,
- a node is considered high frequency if the scales $m \in \{6, 7, 9\}$ contain most of the energy.

For both signals from the previous question (smoothest and least smooth) as well as the first available timestamp, apply this procedure and display on the map the result (one colour per class).

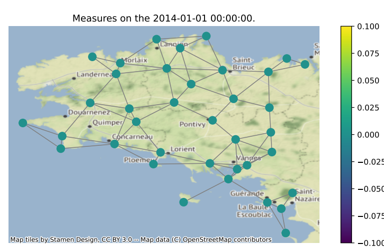
Answer 5



(a) Least smooth signal



(b) Smoothest signal



(c) First available timestamp

Figure 2: Classification of nodes into low/medium/high frequency

Question 6

Display the average temperature and for each timestamp, adapt the marker colour to the majority class present in the graph (see notebook for more details).

Answer 6

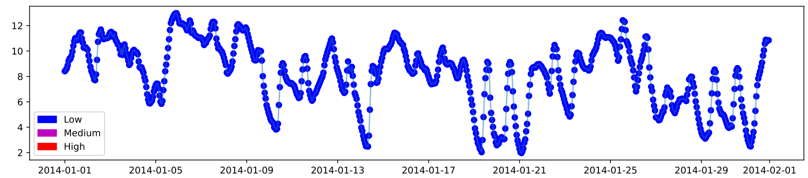


Figure 3: Average temperature. Markers' colours depend on the majority class.

We can see that all, for all the timestamps, the most dominant class is the low-frequency one.

Question 7

The previous graph G only uses spatial information. To take into account the temporal dynamic, we construct a larger graph H as follows: a node is now *a station at a particular time* and is connected to neighbouring stations (with respect to G) and to itself at the previous timestamp and the following timestamp. Notice that the new spatio-temporal graph H is the Cartesian product of the spatial graph G and the temporal graph G' (which is simply a line graph, without loop).

- Express the Laplacian of H using the Laplacian of G and G' (use Kronecker products).
- Express the eigenvalues and eigenvectors of the Laplacian of H using the eigenvalues and eigenvectors of the Laplacian of G and G'.
- Compute the wavelet transform of the temperature signal.
- Classify nodes into low/medium/high frequency and display the same figure as in the previous question.

Answer 7

- Let us denote by n the number of vertices of G and by m the number of vertices of G'. Then, using the Kronecker product, as $H = G \times G'$, we can express the Laplacian $L(H)$ of H :

$$\boxed{L(H) = L(G) \otimes I_m + I_n \otimes L(G')}$$

- Let λ be a eigenvalue of $L(G)$ associated to the eigenvector X and let μ be a eigenvalue of $L(G')$ associated to the eigenvector Y. Then, we have

$$\begin{aligned} L(H)(X \otimes Y) &= (L(G) \otimes I_n + L(G') \otimes I_m)(X \otimes Y) \\ &= L(G)X \otimes I_m Y + I_n X \otimes L(G')Y \\ &= \lambda X \otimes Y + \mu_j X \otimes Y \\ &= (\lambda + \mu_j)(X \otimes Y) \end{aligned}$$

So, we have that the eigenvalues of $L(H)$ are

$$\mathrm{Sp}(L(H)) = \{\lambda + \mu,\ \lambda \in \mathrm{Sp}(L(G)),\ \mu \in \mathrm{Sp}(L(G'))\}$$

and the associated eigenvector of $\lambda + \mu \in \mathrm{Sp}(L(H))$ is $X \otimes Y$ where $L(G)X = \lambda X$ and $L(G')Y = \mu Y$.

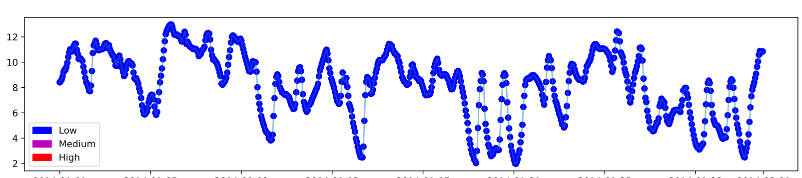


Figure 4: Average temperature. Markers' colours depend on the majority class.

We can therefore see that, when we build the larger graph H to take into account the temporal dynamic, we still have that the msot dominant class id the low-frequency one.