Time series representations: a state-of-the-art

Cyrille Ponchateau

Laboratoire LIAS
cyrille.ponchateau@ensma.fr
www.lias-lab.fr
ISAE - ENSMA

11 Juillet 2016

Content

Content

Time Series

Time series are a temporal data class consisting in a list of values ordered chronologically (Fu 2011).

Time Series

Time series are a temporal data class consisting in a list of values ordered chronologically (Fu 2011).

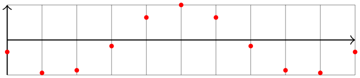


Figure : Drawn time series example

Time Series

Time series are a temporal data class consisting in a list of values ordered chronologically (Fu 2011).

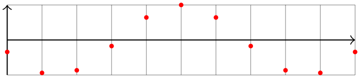


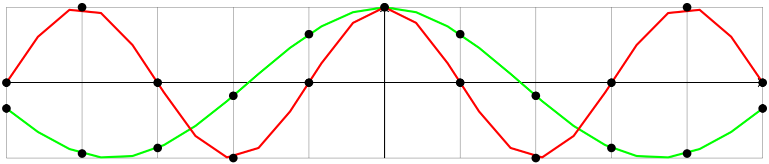
Figure : Drawn time series example

Timestamp	Valeurs
0.0000000e+00	0.0000000e+00
1.0000000e+00	9.4105346e-02
2.0000000e+00	1.8452077e-01
3.0000000e+00	2.7139095e-01
4.0000000e+00	3.5485491e-01
5.0000000e+00	4.3504619e-01
6.0000000e+00	5.1209313e-01
7.0000000e+00	5.8611902e-01
...	...

Table : Example of time series

- Used in many scientific fields.
- Needs efficient analysis method, but...
- Real values => no exact matching.
- Time series quality can be degraded by noise, outliers.
- The analysis must compare time series as a whole.

- Used in many scientific fields.
- Needs efficient analysis method, but...
- Real values => no exact matching.
- Time series quality can be degraded by noise, outliers.
- The analysis must compare time series as a whole.



Dimensionality curse

In addition to the previously mentioned difficulties, time series suffer from the so called "dimensionality curse" problem:

- High-dimensional (one billion items)
- => Raw data processing is costly ;
- => "dimensionality reduction".

Content

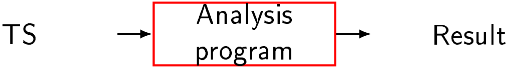


Figure : Naive processing chain

Dimensionality reduction: general idea

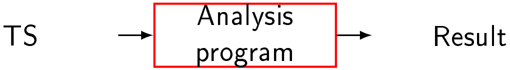


Figure : Naive processing chain

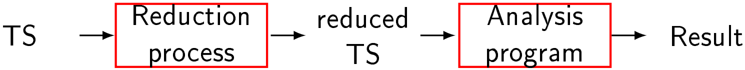


Figure : Dimensionality reduction processing chain

Time series representations categories

The different representations of time series are classified into three categories (Fu 2011):

- *temporal representations*: keep a temporal point of view on the data ;
- *spectral representations*: project it into the frequency domain ;
- *other representations*: apply some other transformation that do not fit into the two first categories.

Simple sampling

The most simple method, consist in sampling the series at fixed length (Åström 1969).

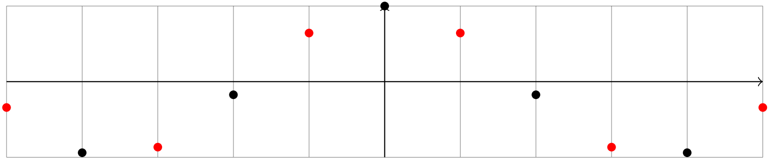


Figure : Sampling example

Extrema extraction

The method consist in extracting all the local extrema of the series. Then the extrema are sorted in terms of importance, in order to eliminate some more points using an importance threshold value (Fink and Gandhi 2011).

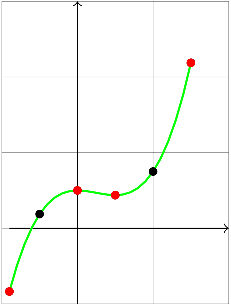


Figure : All extrema extraction

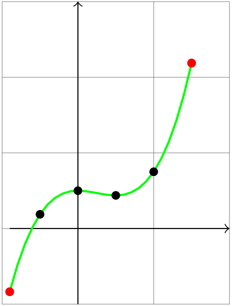


Figure : Extrema sorting

Bit level representation

Points are represented with only one bit equal to one if the value of the point is greater than the mean value of the series, zero otherwise (Bagnall et al. 2006; Ratanamahatana et al. 2005).

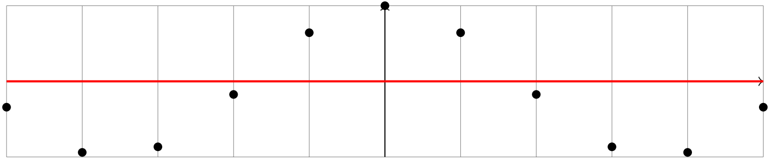


Figure : Bit level representation

Summary of the method seen so far and some more:

- Sampling (Åström 1969) ;
- Extrema extraction (Fink and Gandhi 2011) ;
- Control points identification (Man and Wong 2001) ;
- Landmark model (Perng et al. 2000) ;
- Salient points identification (Perceptually Important Points) (Chung et al. 2001) ;
- Bit level representation (Bagnall et al. 2006; Ratanamahatana et al. 2005).

Discrete Fourier Transform

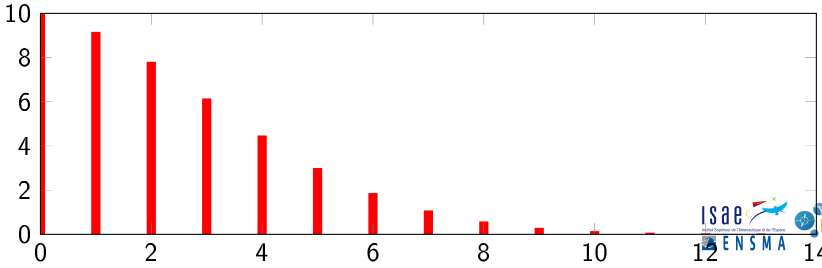
Any continuous periodic function can be transformed into a combination of sines and cosines. The Discrete Fourier Transform maps a discrete periodic sequence f to a discrete sequence of coefficients F, representing the Fourier transform of the sequence (Shatkay 1995; Agrawal, Faloutsos, and Swani 1993). With N equal to the size of f:

$$f[t] = \frac{1}{N} \sum_{j=0}^{N-1} \boxed{F[j]} e^{\frac{2\pi itj}{N}} \tag{1}$$

Discrete Fourier Transform

Any continuous periodic function can be transformed into a combination of sines and cosines. The Discrete Fourier Transform maps a discrete periodic sequence f to a discrete sequence of coefficients F, representing the Fourier transform of the sequence (Shatkay 1995; Agrawal, Faloutsos, and Swani 1993). With N equal to the size of f:

$$f[t] = \frac{1}{N}\sum_{j=0}^{N-1} \boxed{F[j]} e^{\frac{2\pi itj}{N}} \quad (1)$$



Discrete Wavelet Transform

Any continuous function can be decomposed on a basis of wavelets functions. Wavelets are local functions, allowing more accuracy on local details and multiple resolutions (Sidney, Gopinath, and Guo 1997).

Figure : Wavelet example

- Hidden Markov Model (HMM) (Azzouzi and Nabney 1998) ;
- Reference time series decomposition (Kriegel et al. 2008) ;
- AutoRegressive Moving Average (ARMA) (Shumway and Stoffer 2015) ;
- Singular Value Decomposition (SVD) (Korn, Jagadish, and Faloutsos 1997) ;
- SVD with Deltas (SVDD) (Korn, Jagadish, and Faloutsos 1997) ;
- Dictionary Compression Score (DCS) (Lang, Morse, and Patel 2010).

Content

Segment

A segment is a subsequence of a time series, also called time window.

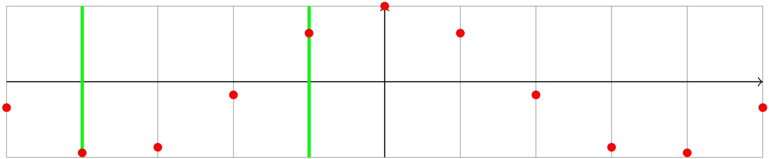


Figure : Segment example

Segmentation

A segmentation process consist in dividing a time series into a set of non-overlapping segments. Inside a segment, the data can be represented by one specific value or a continuous function (generally straight lines) (Lovrić, Milanović, and Stamenković 2014).

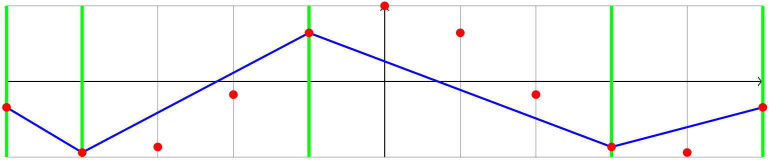


Figure : Segmentation example

Segment data representation

There is different ways to represent the data within a segment:

- Piecewise Aggregate Approximation (PAA) (Keogh et al. 2000) ;
- Adaptive Piecewise Constant Approximation (APCA) (Chakrabarti et al. 2002) ;
- Piecewise Linear Segmentation, with two ways to define the lines (Keogh et al. 2004):
  - Linear interpolation ;
  - Linear regression.
- It is also possible to use polynomials ;
- Or any continuous and differentiable function (Shatkay and Zdonik 1996).

Segmented Sum of Variation

The Segmented Sum of Variation consist in computing the difference between each consecutive value and sum them (Lee, Kwon, and Lee 2003).

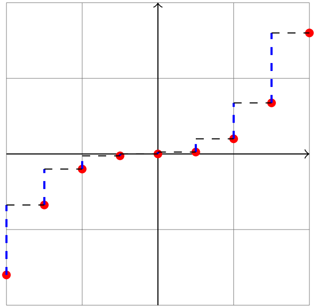


Figure : SSV example

Segment data representation

There is different ways to represent the data within a segment:

- Piecewise Aggregate Approximation (PAA) (Keogh et al. 2000) ;
- Adaptive Piecewise Constant Approximation (APCA) (Chakrabarti et al. 2002) ;
- Piecewise Linear Segmentation, with two ways to define the lines (Keogh et al. 2004):
  - Linear interpolation ;
  - Linear regression.
- It is also possible to use polynomials ;
- Or any continuous and differentiable function (Shatkay and Zdonik 1996).

Segmentation approaches

Segmentation algorithms can be classified into three classes: Top-Down, Bottom-Up and Sliding-Window.

Top-Down

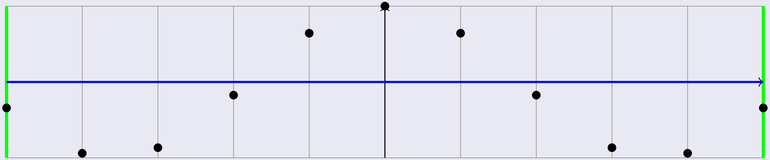


Figure : Top-Down approach example

Segmentation approaches

Segmentation algorithms can be classified into three classes: Top-Down, Bottom-Up and Sliding-Window.

Top-Down

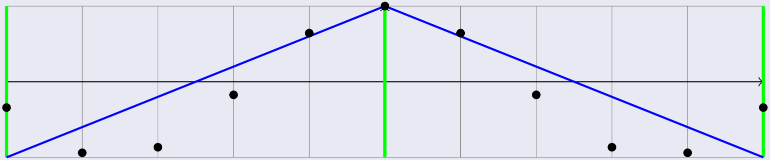


Figure : Top-Down approach example

Segmentation approaches

Segmentation algorithms can be classified into three classes: Top-Down, Bottom-Up and Sliding-Window.

Bottom-Up

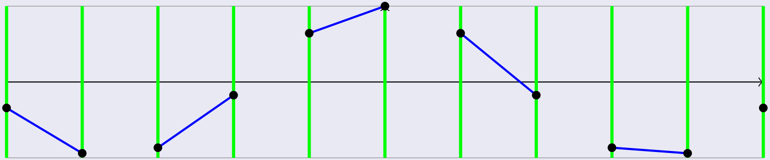


Figure : Bottom-Up approach example

Segmentation approaches

Segmentation algorithms can be classified into three classes: Top-Down, Bottom-Up and Sliding-Window.

Sliding-Window

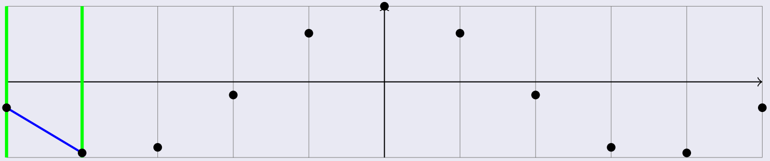


Figure : Sliding-Window approach example

Segmentation approaches

Segmentation algorithms can be classified into three classes: Top-Down, Bottom-Up and Sliding-Window.

Sliding-Window

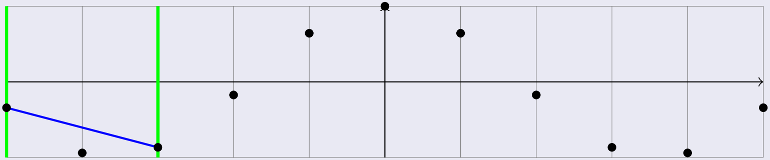


Figure : Sliding-Window approach example

Segmentation approaches

Segmentation algorithms can be classified into three classes: Top-Down, Bottom-Up and Sliding-Window.

Sliding-Window

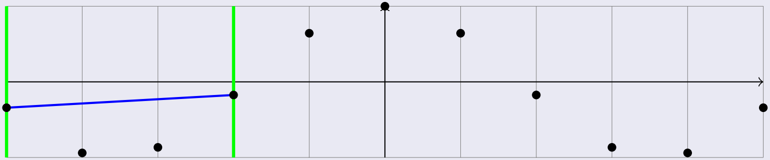


Figure : Sliding-Window approach example

Segmentation approaches

Segmentation algorithms can be classified into three classes: Top-Down, Bottom-Up and Sliding-Window.

Sliding-Window

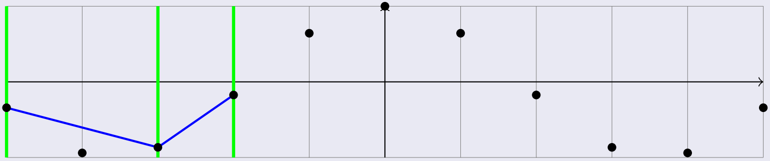


Figure : Sliding-Window approach example

- In most cases, sliding-window gives the worst performances ;
- In some cases, top-down gives slightly better results than bottom-up ;
- In general, bottom-up performances are significantly better than the other two.

Another remark:

- Top-Down and Bottom-Up are offline algorithms ;
- Sliding-Window are online algorithms ;
- SWAB (Sliding-Window And Bottom-Up) online, with performances competing with Bottom-Up (Keogh et al. 2004).

Content

- Lots of different methods, for numerous algorithms and purposes.
- Lots of confusion on the superiority of a solution compared to others.
- DFT >< DWT ; PAA >< APCA ...
- What is the best methods ?
  - What is the objective ?
  - Is their any assumptions on the time series ?

Thank you.

Do you have questions ?