Lesson 13 Cluster Analysis for Time Series

Lusine Zilfimian

May 17 (Wednesday), 2023

Contents

- Quiz

Contents

- Quiz
- DTW

DTW vs Euclidean distance

- Euclidean distance has become one of the most commonly used distance measures when working with time series

DTW vs Euclidean distance

- Euclidean distance has become one of the most commonly used distance measures when working with time series
- Due to **linear aligning** of related points of time series, it is very sensitive to distortions along the time axis

DTW vs Euclidean distance

- Euclidean distance has become one of the most commonly used distance measures when working with time series
- Due to **linear aligning** of related points of time series, it is very sensitive to distortions along the time axis
- DTW finds optimal **non-linear alignment**

DTW

- Dynamic time warping is a popular technique for comparing time series, providing both a **distance measure** and the **warping**.

DTW

- Dynamic time warping is a popular technique for comparing time series, providing both a **distance measure** and the **warping**.
- The main purpose of Dynamic Time Warping (DTW) in time series clustering is to measure the **similarity** between two time series, even if they have different lengths or exhibit time distortions.

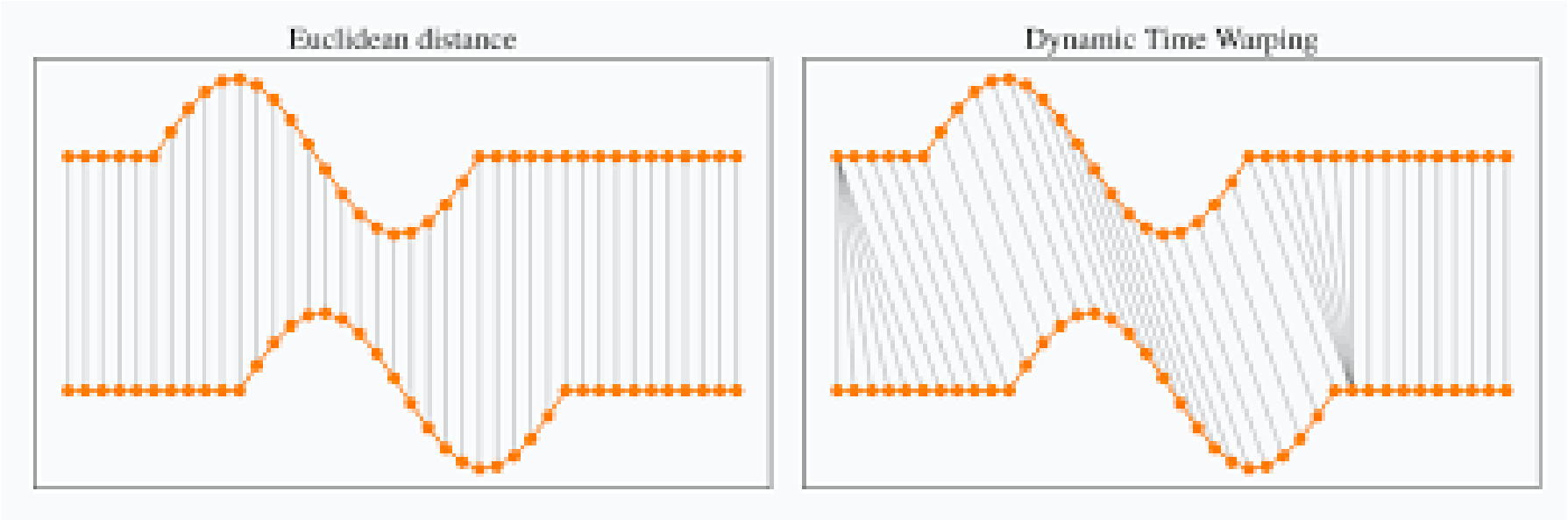
DTW

- Dynamic time warping is a popular technique for comparing time series, providing both a **distance measure** and the **warping**.
- The main purpose of Dynamic Time Warping (DTW) in time series clustering is to measure the **similarity** between two time series, even if they have different lengths or exhibit time distortions.
- DTW helps to find the best alignment between the two time series by warping, stretching, or compressing them along the time axis.

DTW

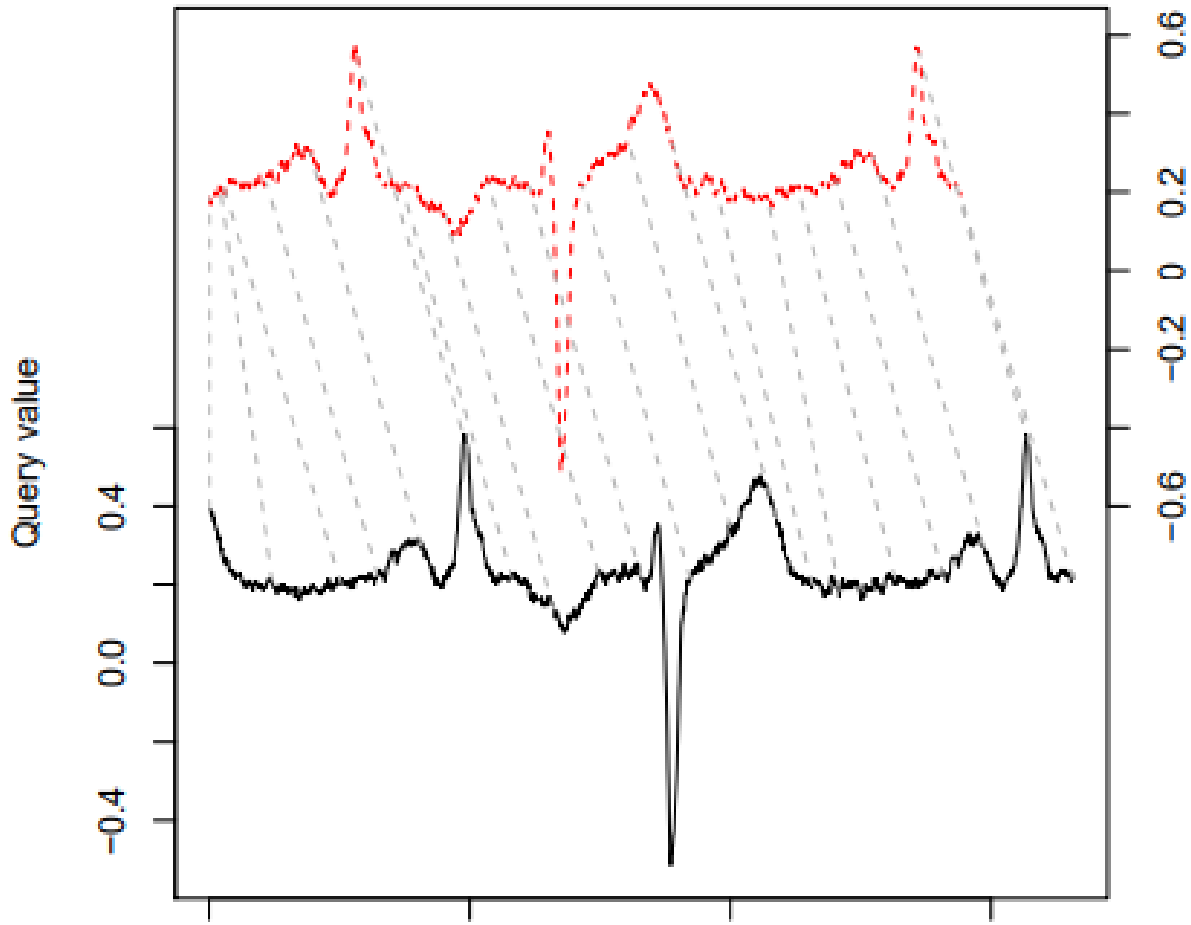
- Dynamic time warping is a popular technique for comparing time series, providing both a **distance measure** and the **warping**.
- The main purpose of Dynamic Time Warping (DTW) in time series clustering is to measure the **similarity** between two time series, even if they have different lengths or exhibit time distortions.
- DTW helps to find the best alignment between the two time series by warping, stretching, or compressing them along the time axis.
- It finds the alignment that minimizes the total distance between corresponding points in the time series.

Idea of DTW



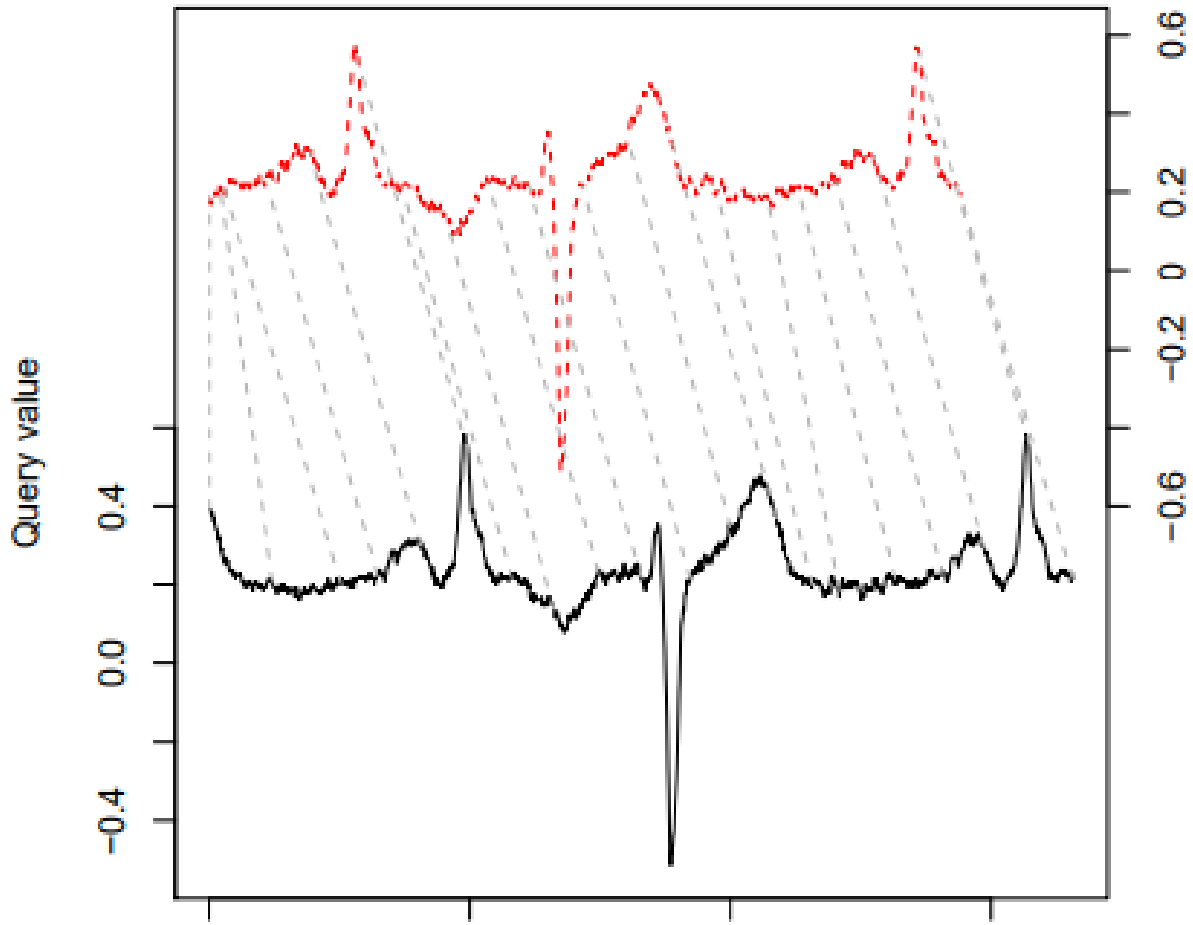
Idea of CA

- Make one time series resembles the other as much as possible.



Idea of CA

- Make one time series resembles the other as much as possible.
- We can stretch or compress them.



Variables/Parameters

- Continuity

Variables/Parameters

- Continuity
- Endpoints

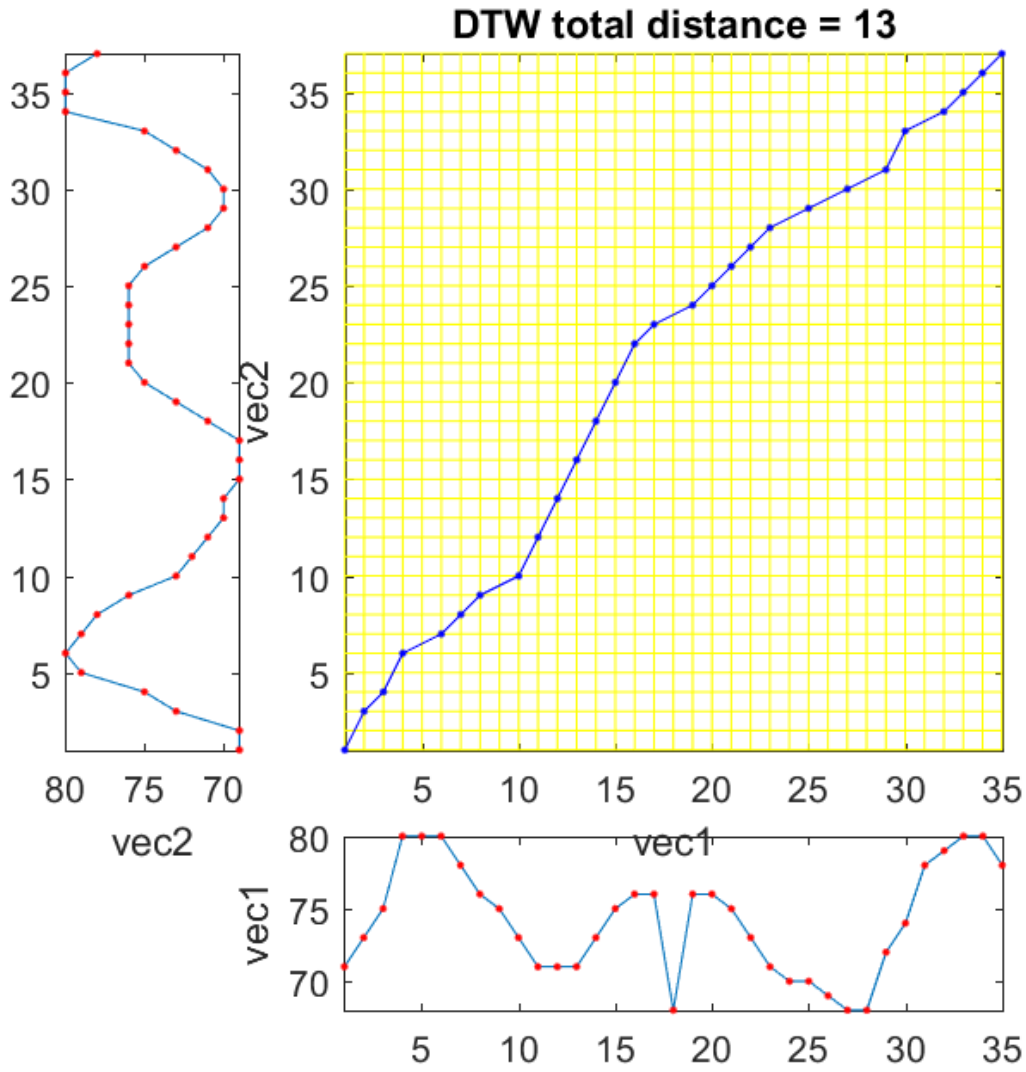
Variables/Parameters

- Continuity
- Endpoints
- Local distance definitions

Variables/Parameters

- Continuity
- Endpoints
- Local distance definitions
- Global constraints

Matrix of correlations and path



Type of step functions

```
dtw(a1, a2)$stepPattern
```

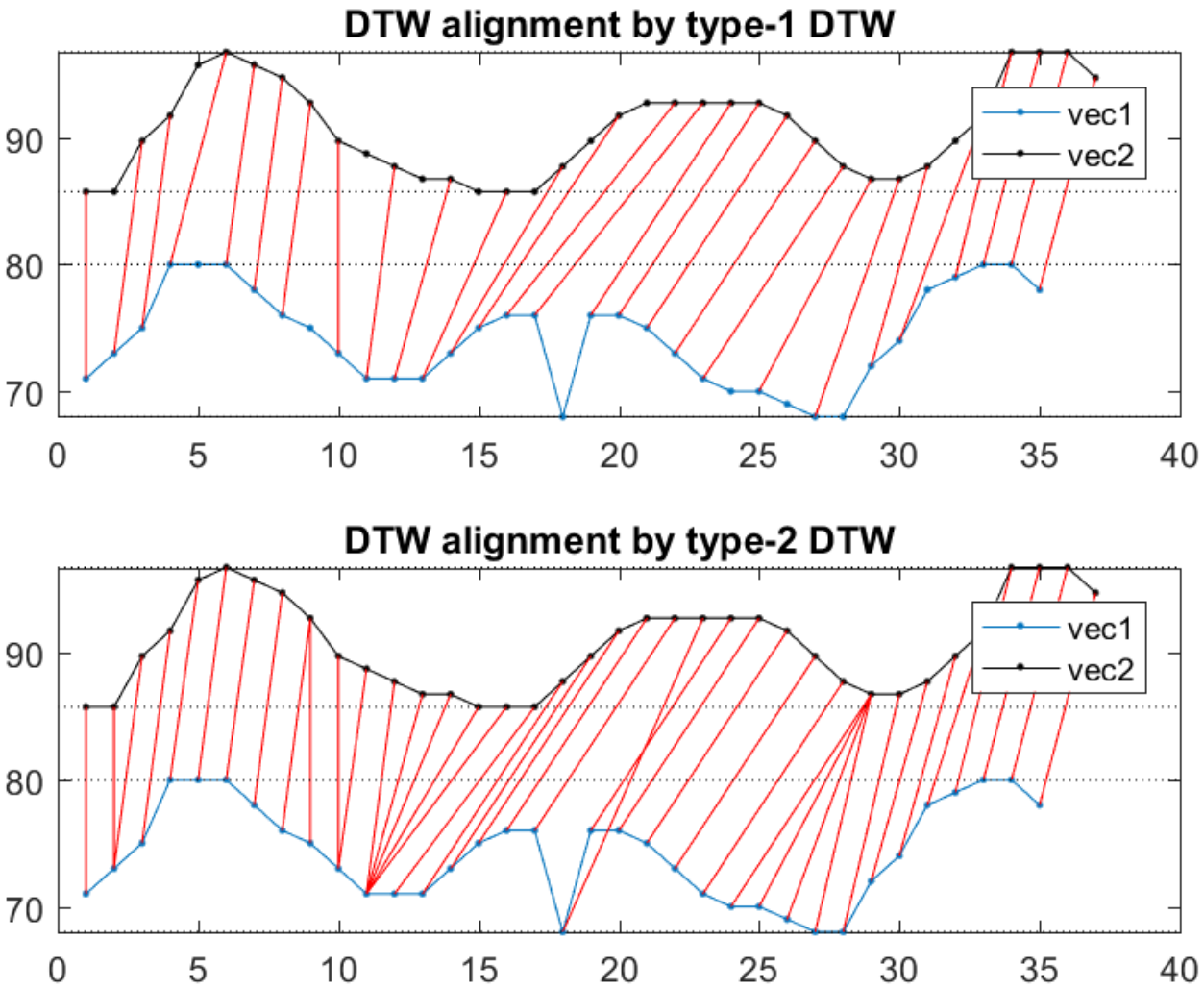
```
## Step pattern recursion:
## g[i,j] = min(
##      g[i-1,j-1] + 2 * d[i  ,j  ] ,
##      g[i  ,j-1] +     d[i  ,j  ] ,
##      g[i-1,j  ] +     d[i  ,j  ] ,
##    )
##
##  Normalization hint: N+M
```

Type of step functions

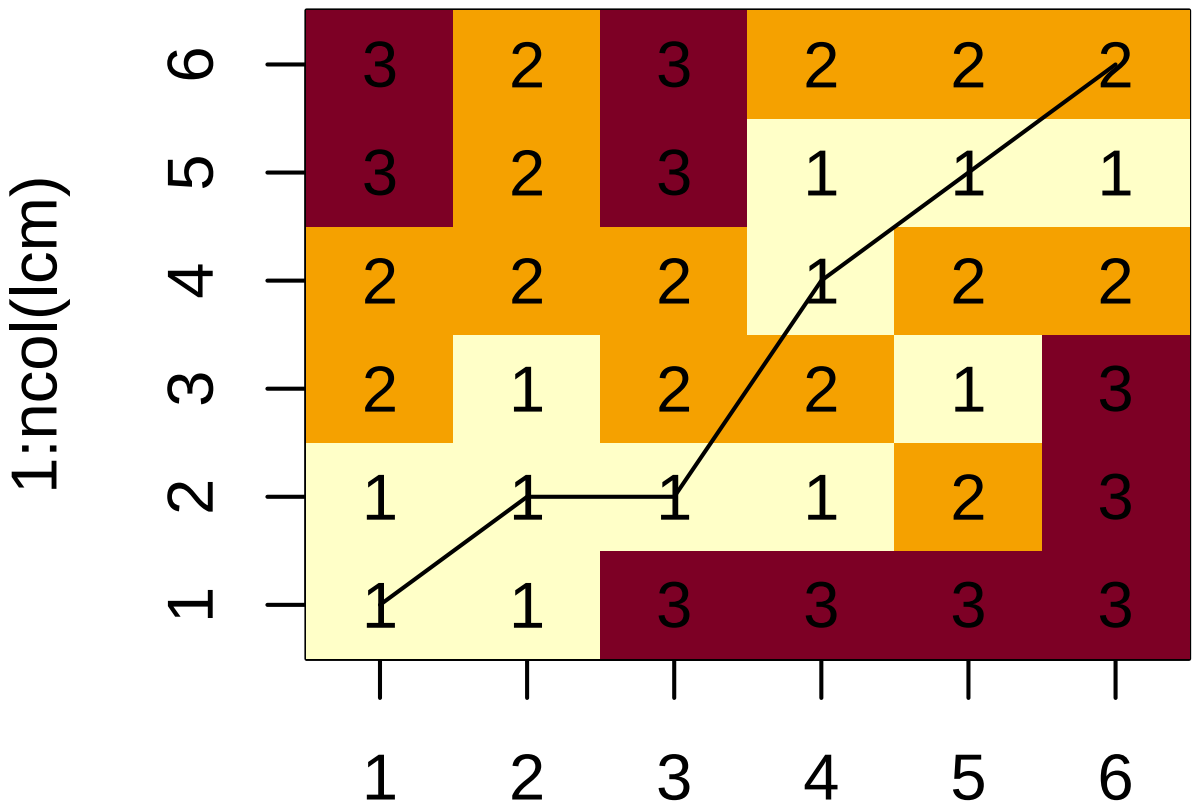
```
dtw(a1, a2, step.pattern = symmetric1)$stepPattern
```

```
## Step pattern recursion:
## g[i,j] = min(
##      g[i-1,j-1] +     d[i  ,j  ] ,
##      g[i  ,j-1] +     d[i  ,j  ] ,
##      g[i-1,j  ] +     d[i  ,j  ] ,
##    )
##
##  Normalization hint: NA
```

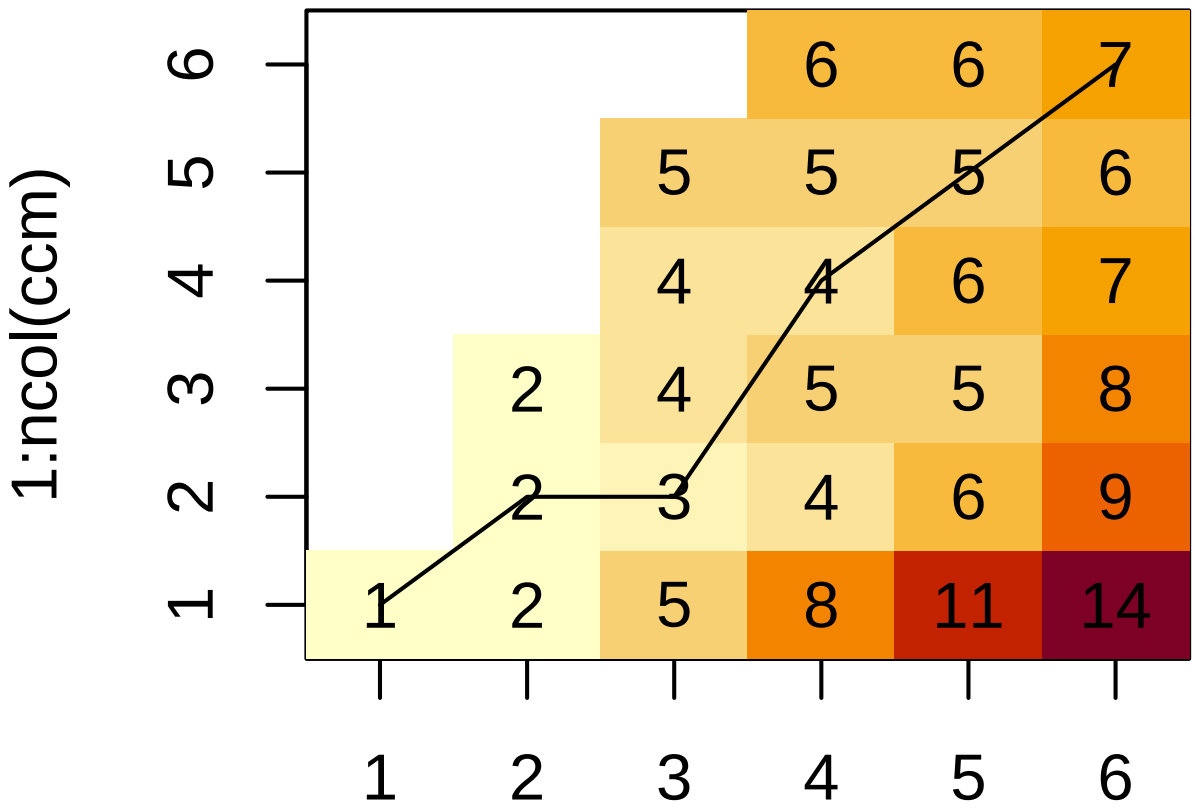
Type of step functions



Local distance matrix



Cumulative distance matrix



Problems

- Finding the optimal alignment can be quite **slow**

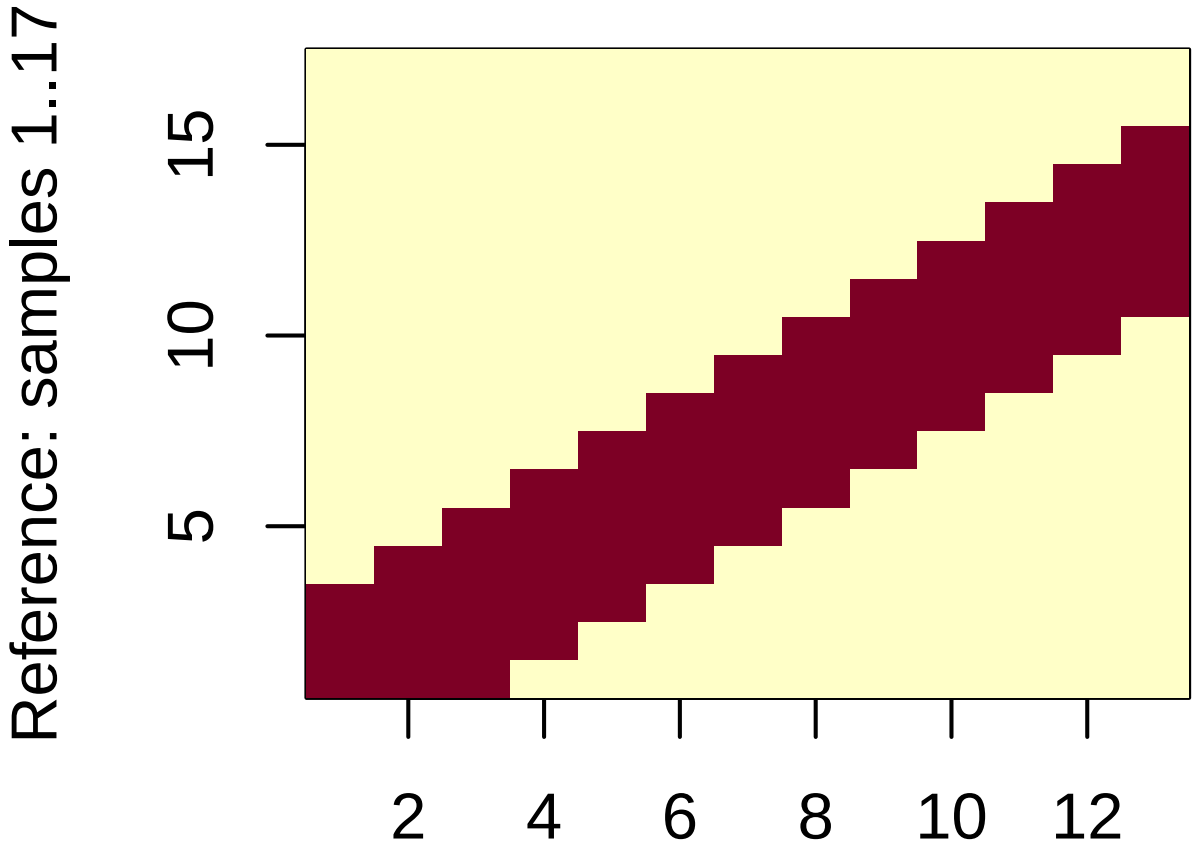
Problems

- Finding the optimal alignment can be quite **slow**
- It can lead to bad alignments where a relatively small part of one time series maps onto a large section of the other one.

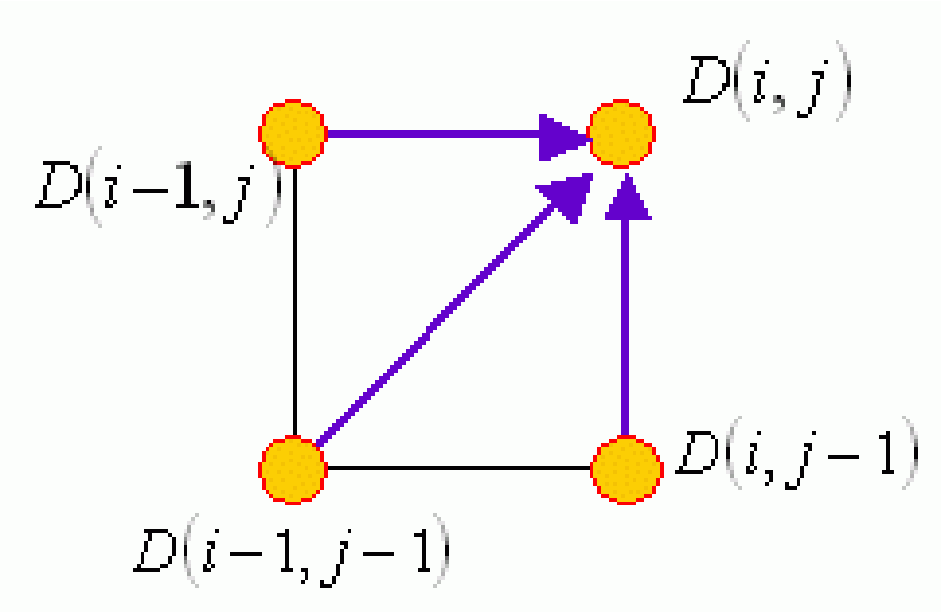
Problems

- Finding the optimal alignment can be quite **slow**
- It can lead to bad alignments where a relatively small part of one time series maps onto a large section of the other one.
- This can be avoided by narrowing the searching window around the diagonal of the warping matrix using global constraints.

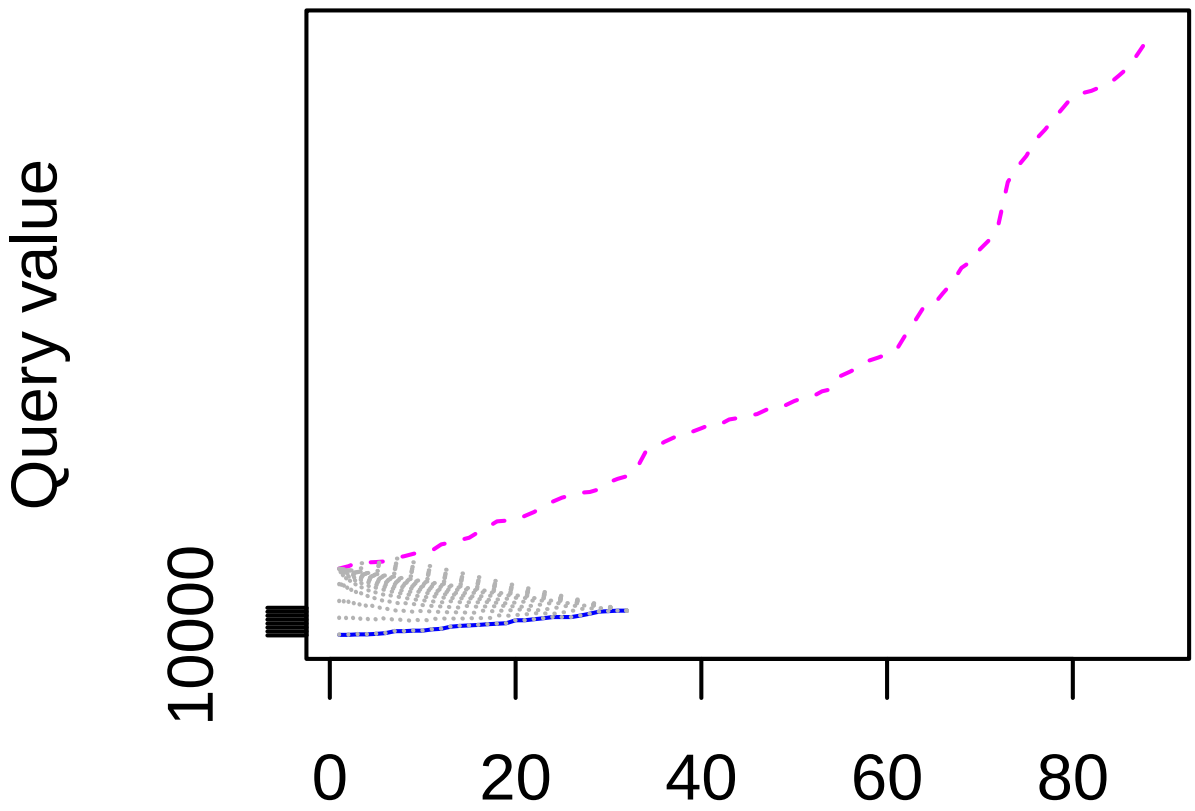
Global Path Constraint: Sakoe Chiba Window



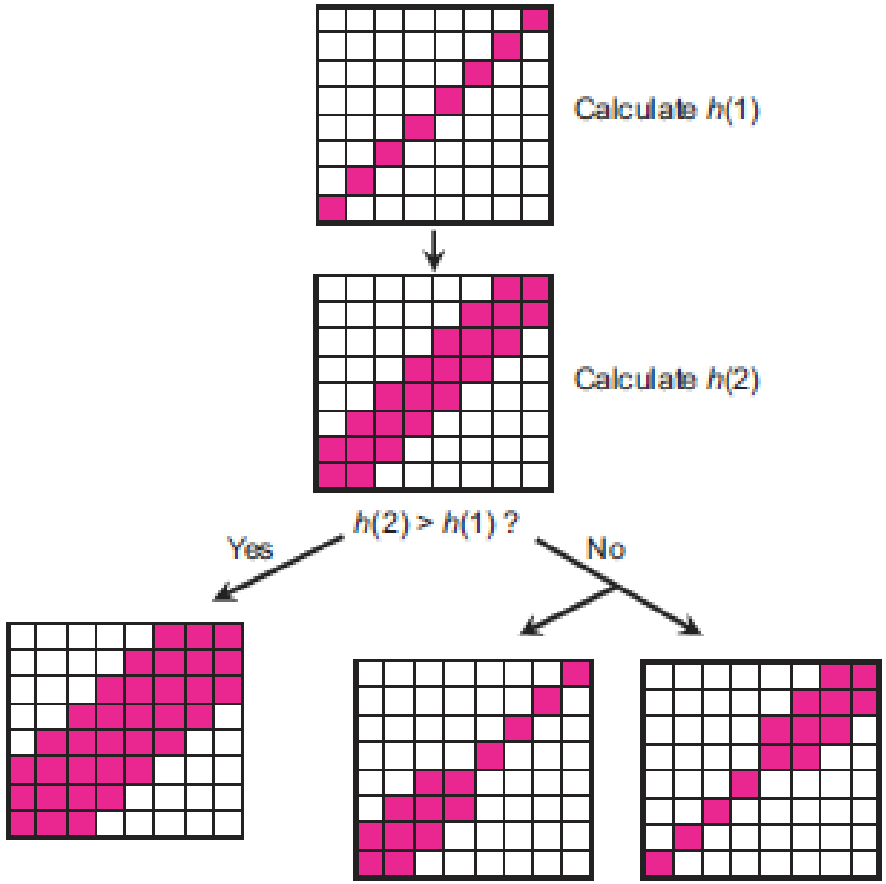
Local Path contraint



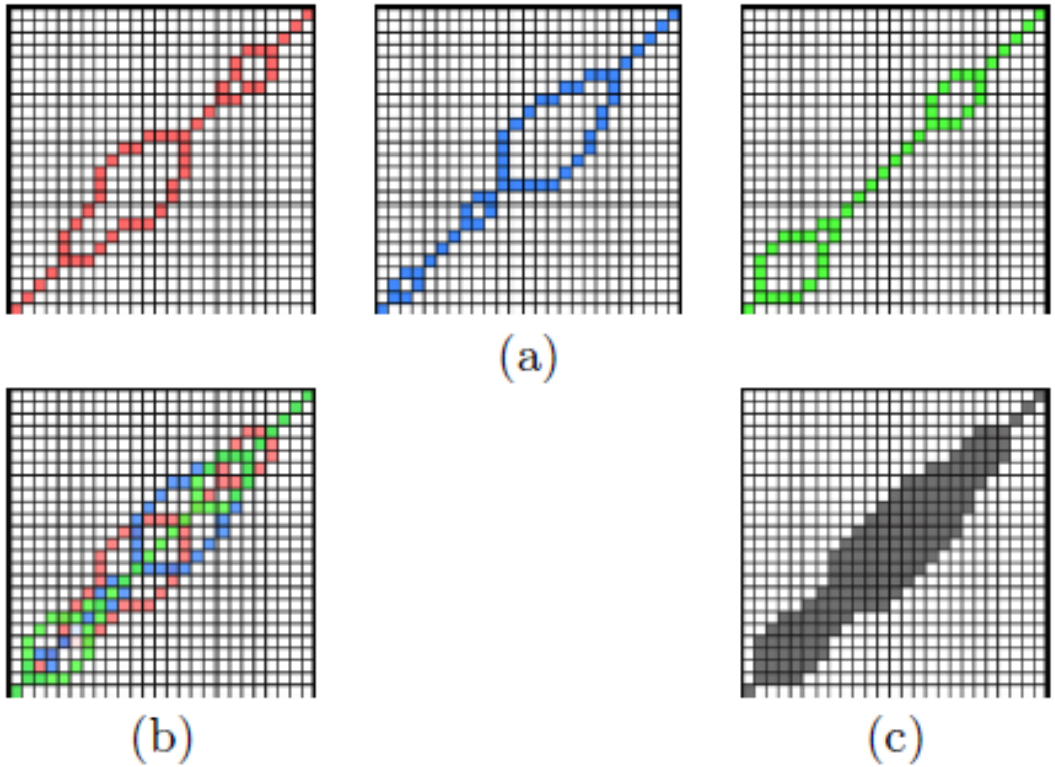
Unconstrained endpoints



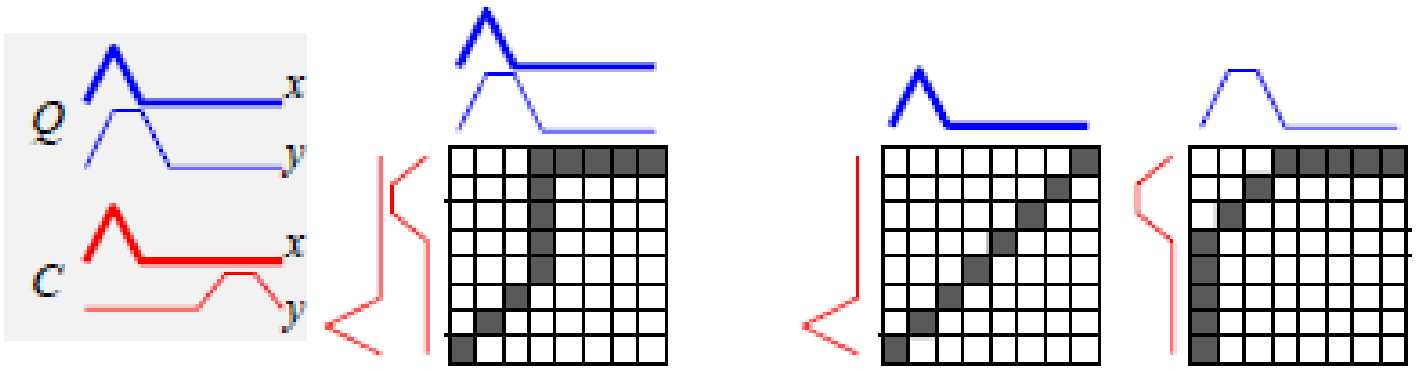
R-K band forward searching algorithm



R-K band forward searching algorithm



Dealing with multivariate time series



And

The end.