DATA MINING 2
Exercises – Time Series

Riccardo Guidotti

a.a. 2019/2020

UNIVERSITÀ DI PISA

Dynamic Time Warping

DTW – Exercise 1

- Given the following input time series:

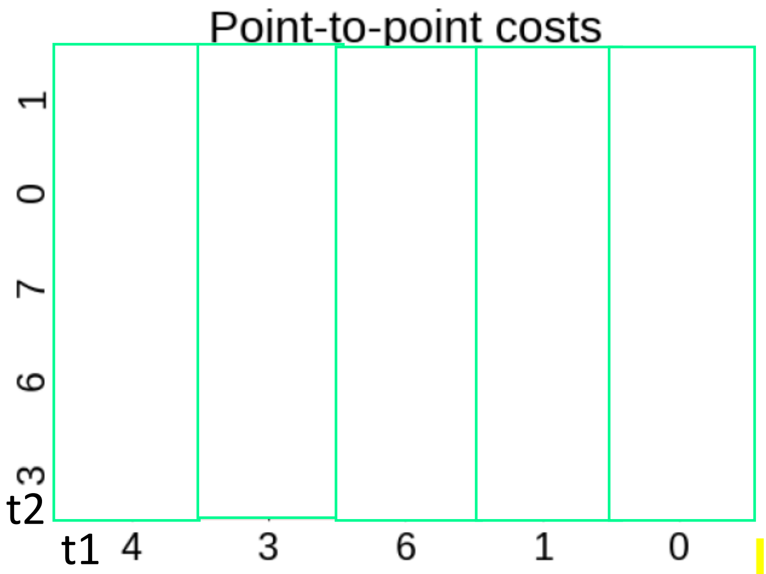
t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A) Compute the distance between “t1” and “t2”, using the DTW with distance between points computed as d(x,y) = |x – y|.
- B) If we repeat the computation of point (A) above, this time with a Sakoe-Chiba band of size r=1, does the result change? Why?
- C) If we compute DTW(T1,T2), where T1 is equal to t1 in reverse order (namely T1=<0,1,6,3,4>) and similarly for T2 (namely T2=<1,0,7,6,3>), is it true that DTW(T1,T2) = DTW(t1,t2)? Discuss the problem without providing any computation.

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)

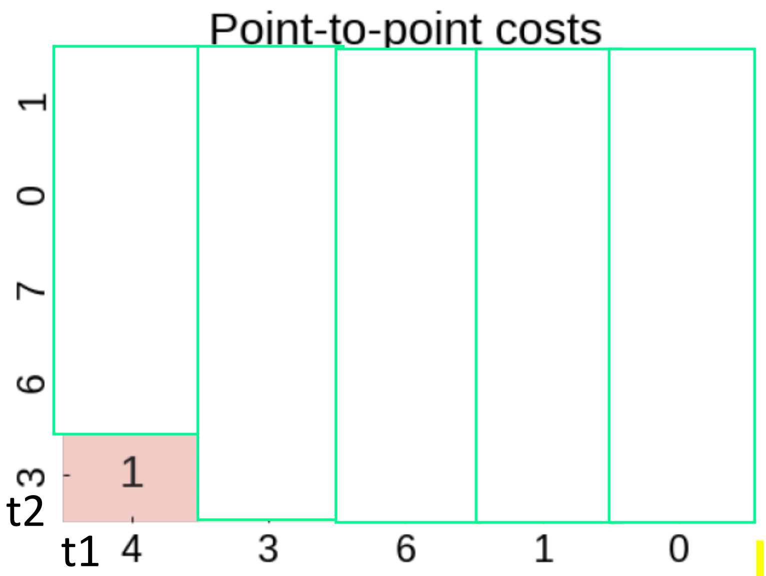


Result: 4

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)

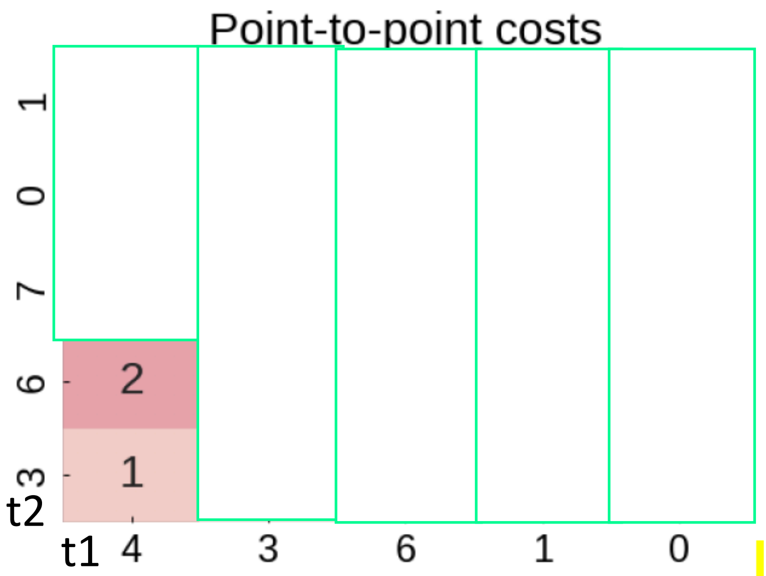


Result: 4

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)

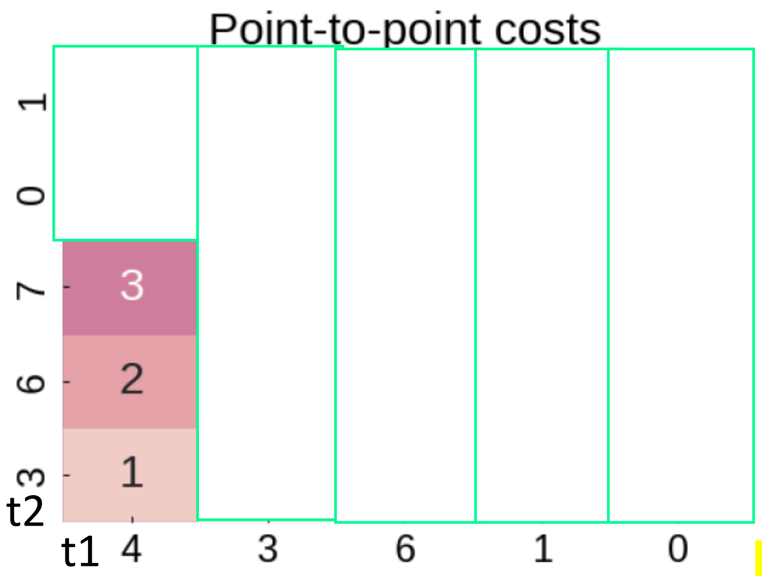


Result: 4

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)

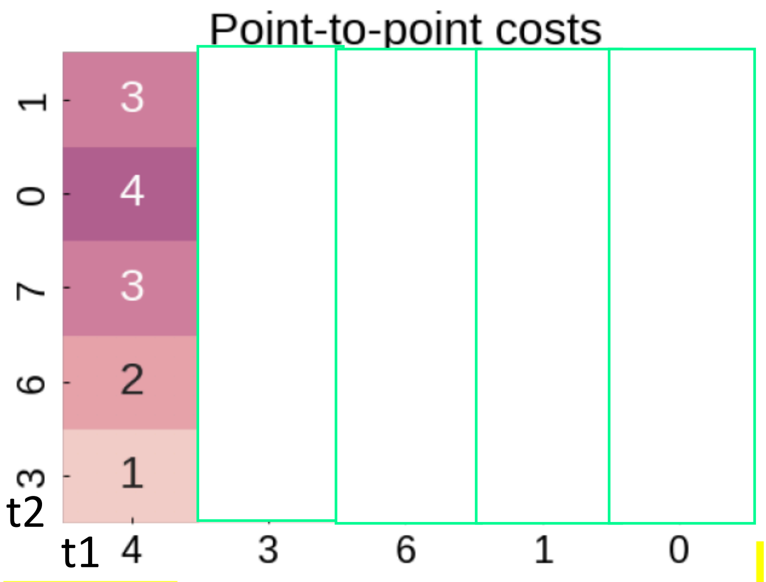


Result: 4

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)

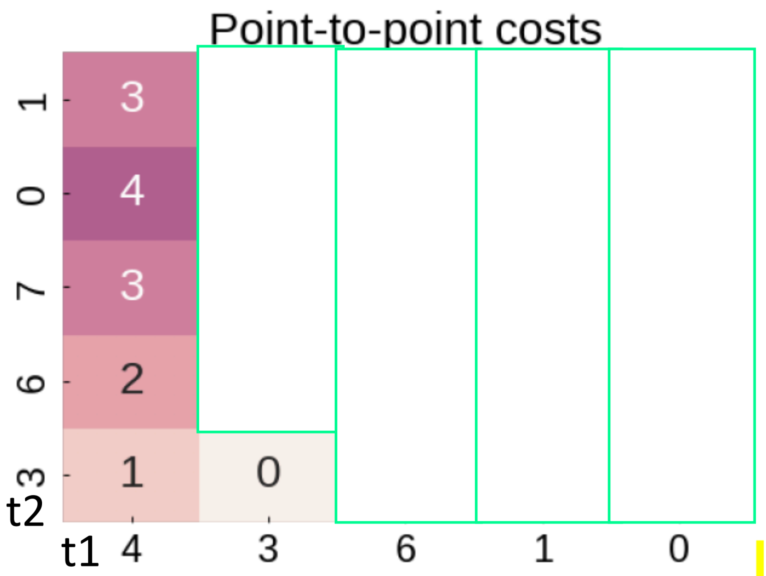


Result: 4

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)

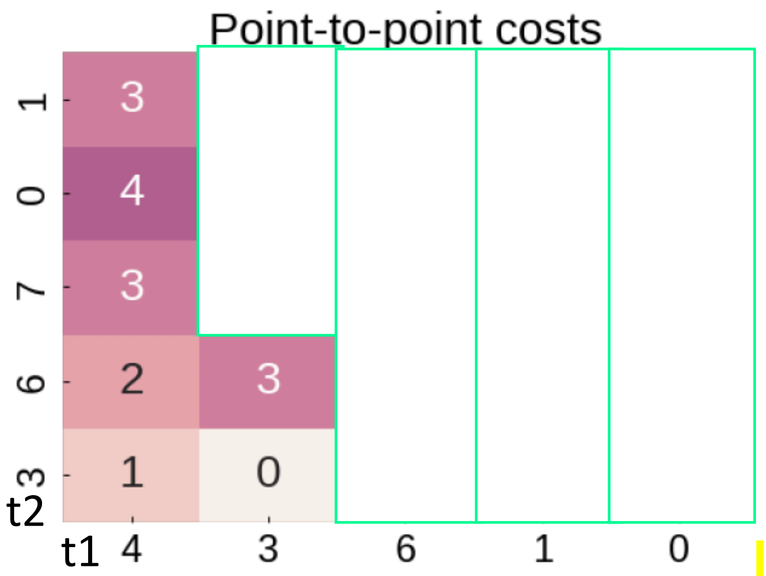


Result: 4

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)

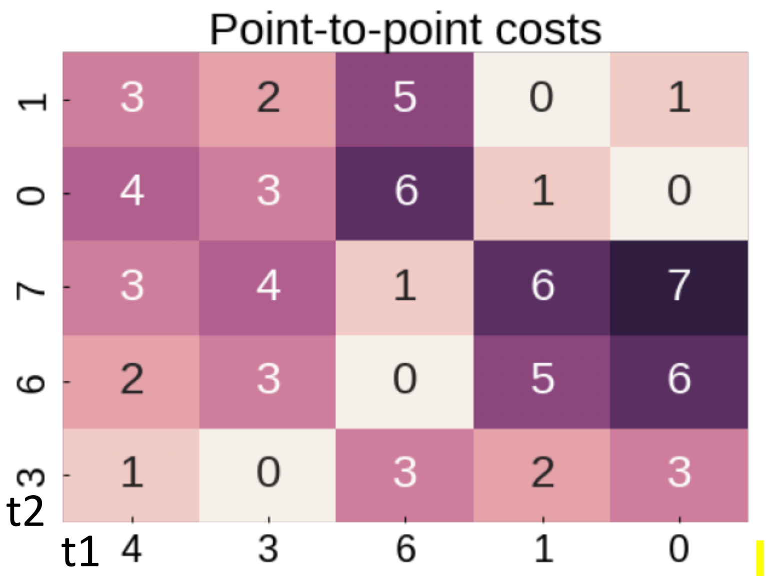


Result: 4

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)

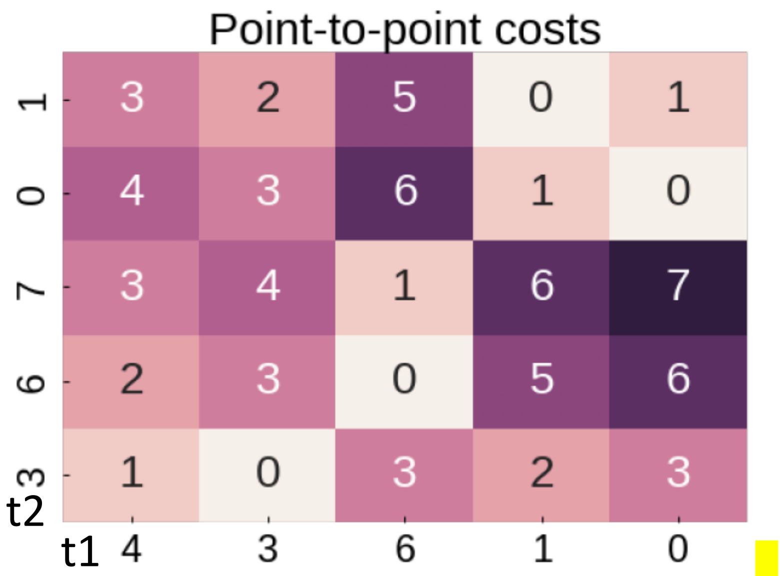


Result: 4

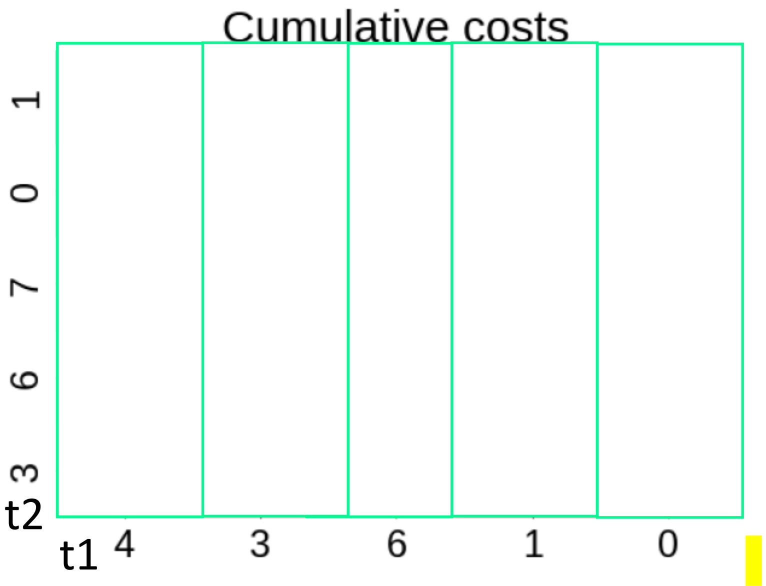
DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)



Result: 4

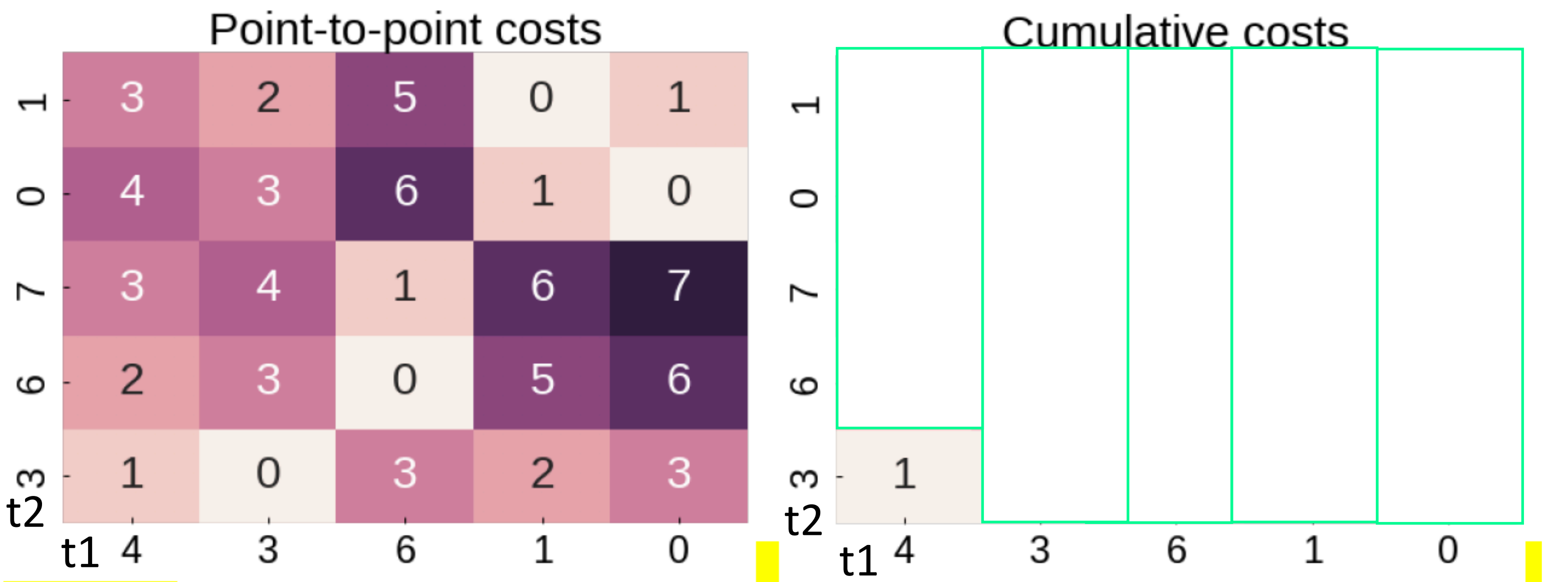


$$\gamma(i,j) = d(q_i,c_j) + \min\{ \gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1) \}$$

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)



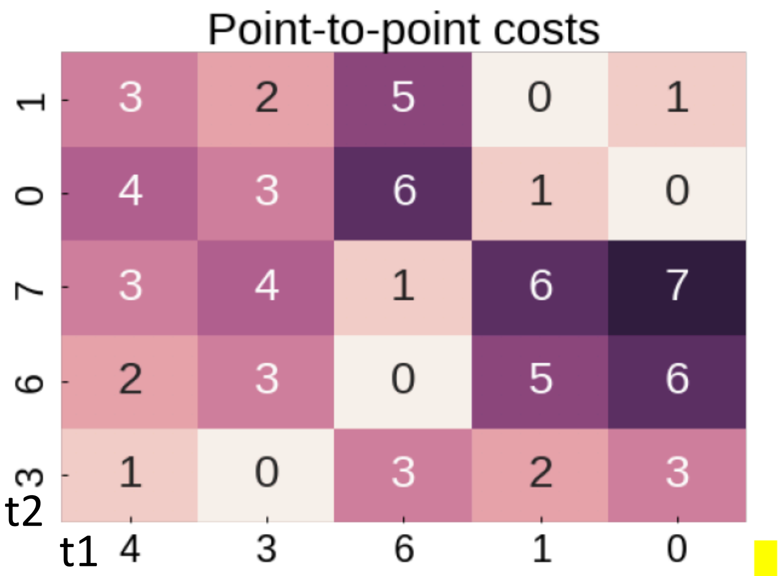
Result: 4

$$\gamma(i,j) = d(q_i,c_j) + \min\{ \gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1) \}$$

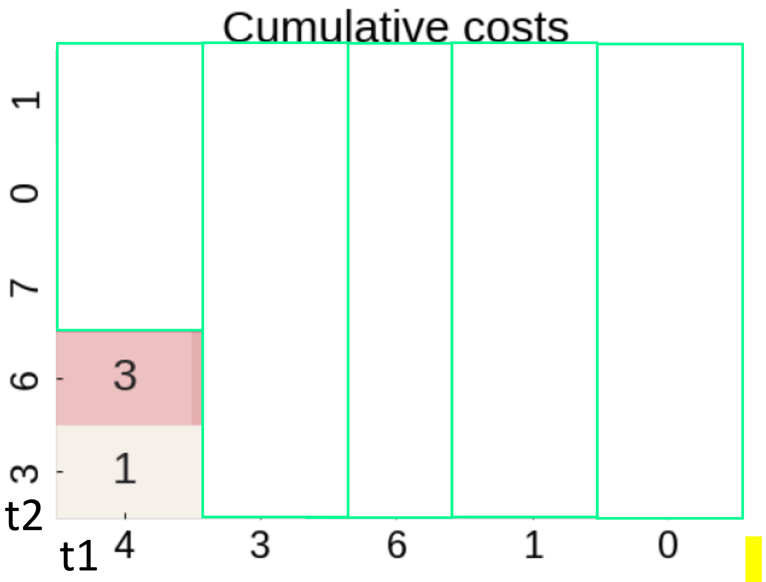
DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)



Result: 4

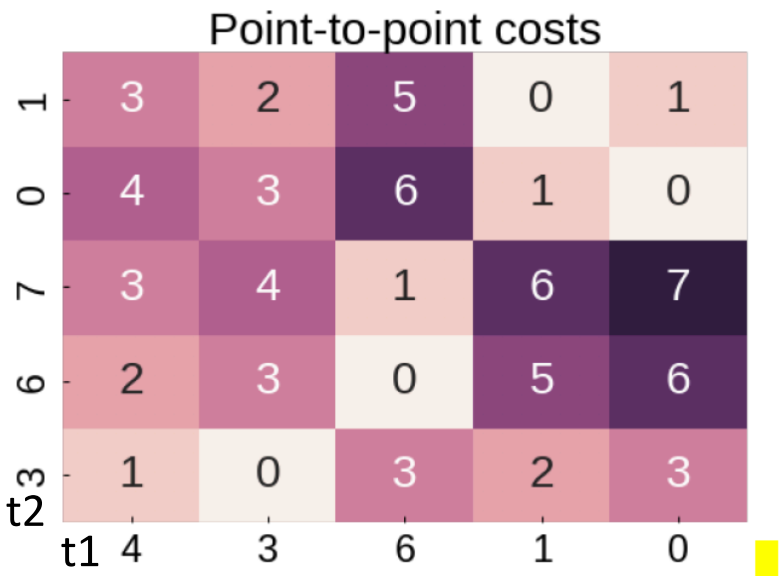


$$\gamma(i,j) = d(q_i,c_j) + \min\{ \gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1) \}$$

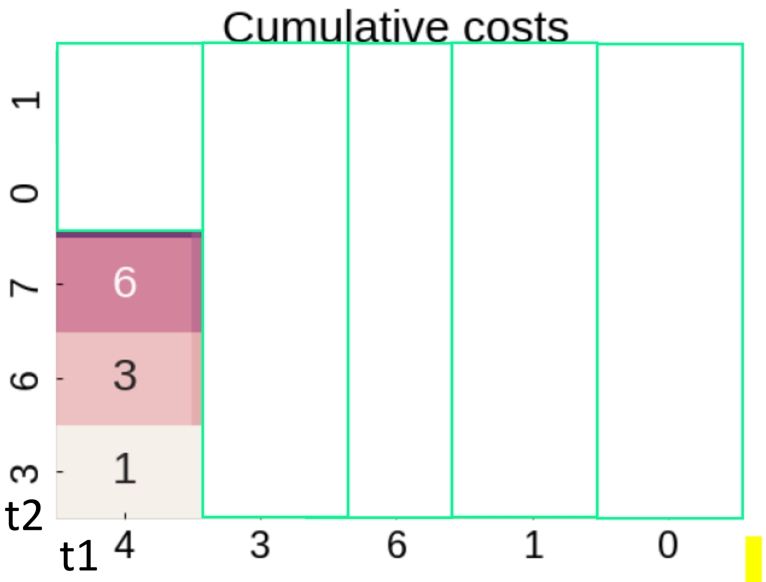
DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)



Result: 4

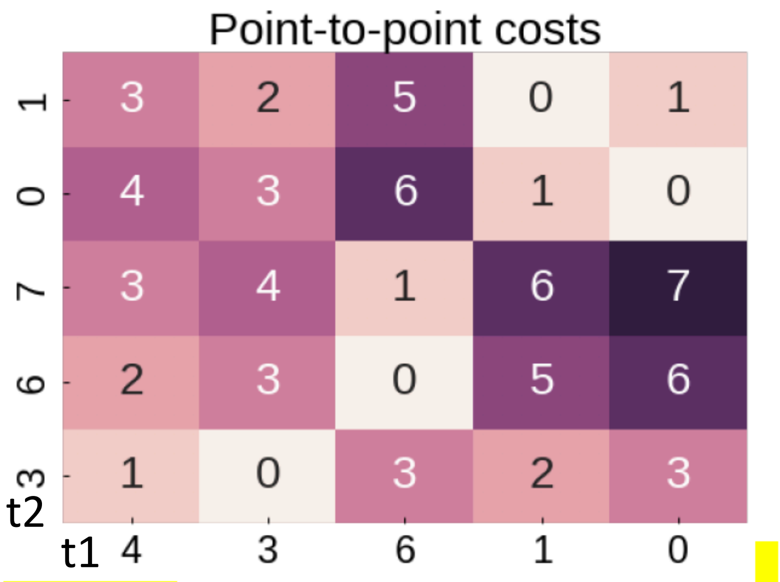


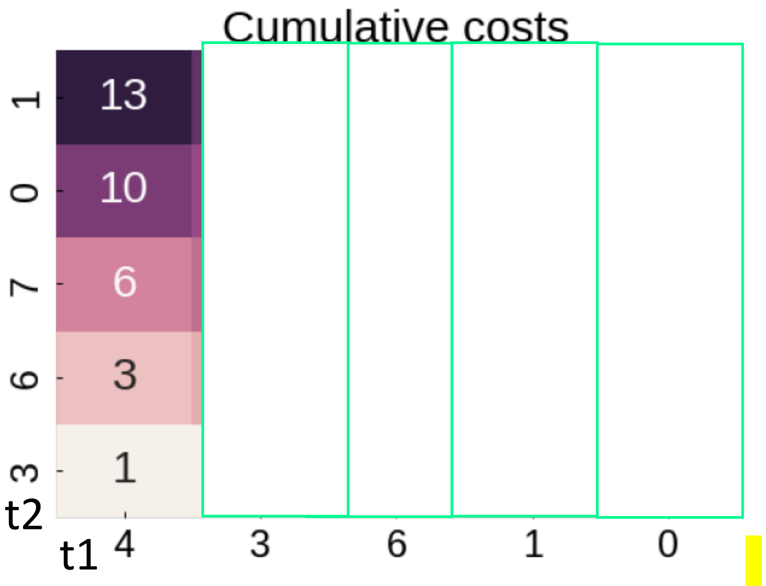
$$\gamma(i,j) = d(q_i,c_j) + \min\{ \gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1) \}$$

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)





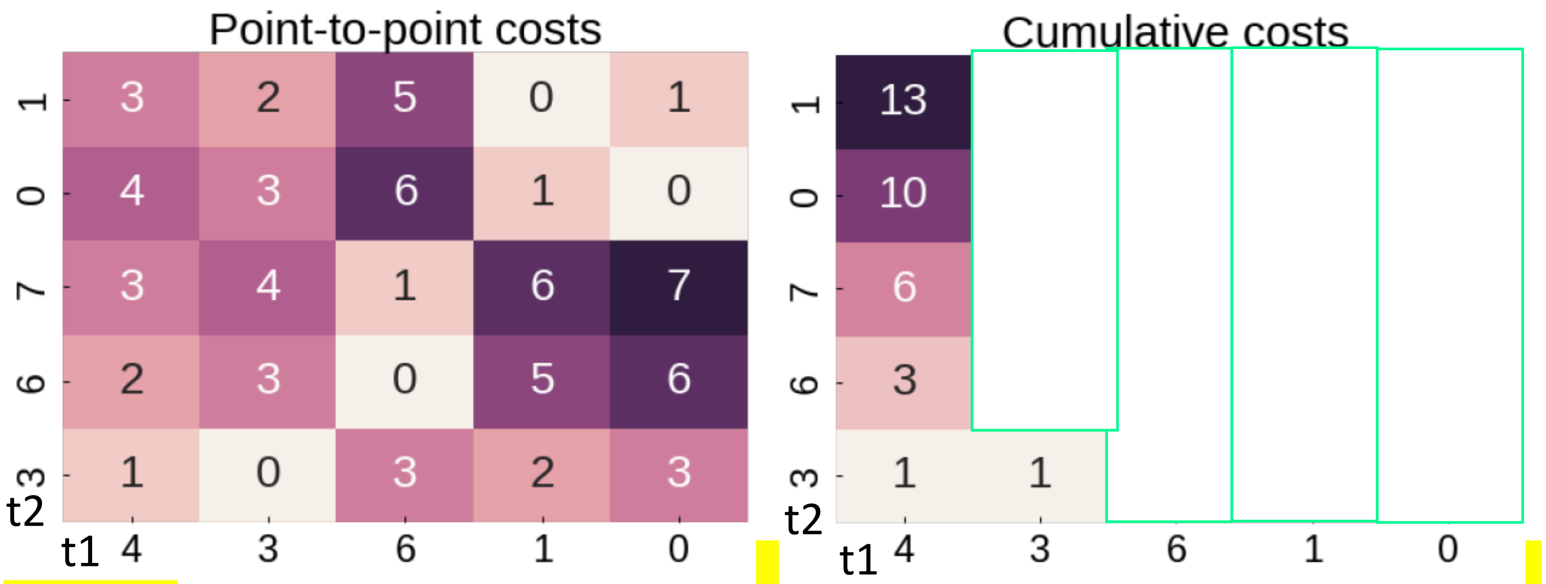
Result: 4

$$\gamma(i,j) = d(q_i,c_j) + \min\{ \gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1) \}$$

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)



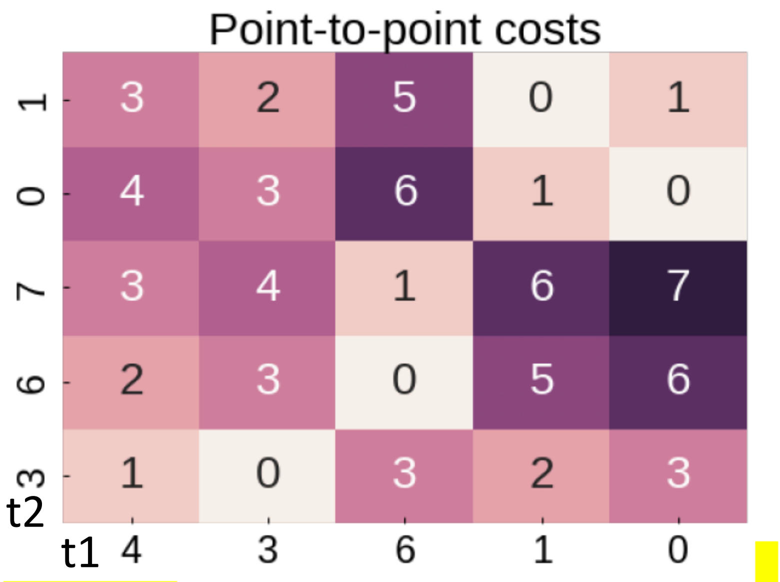
Result: 4

$$\gamma(i,j) = d(q_i, c_j) + \min\{\, \gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1) \,\}$$

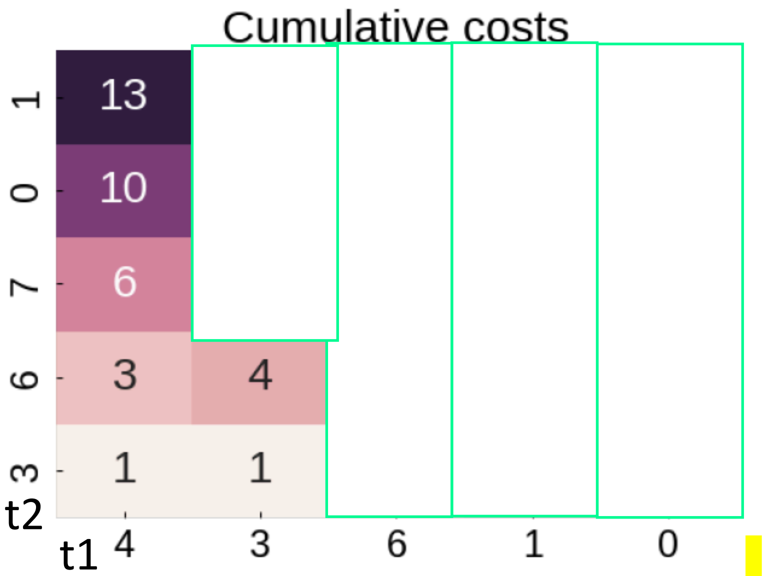
DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)



Result: 4

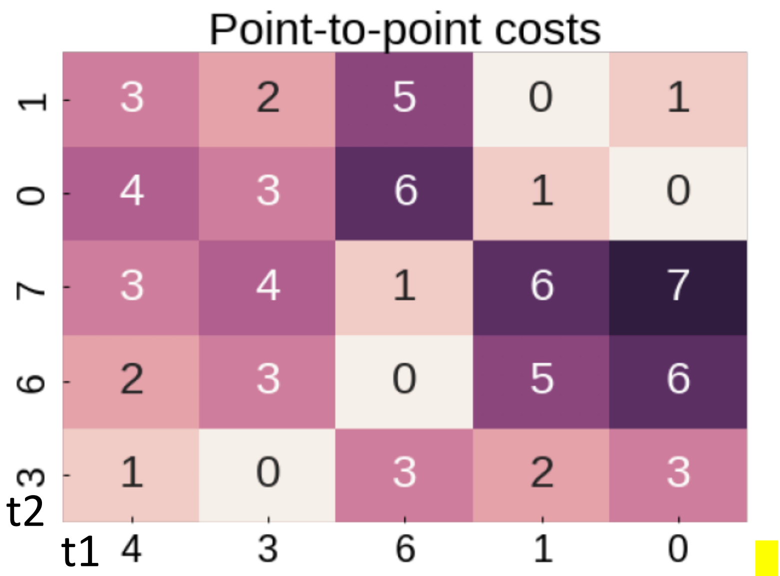


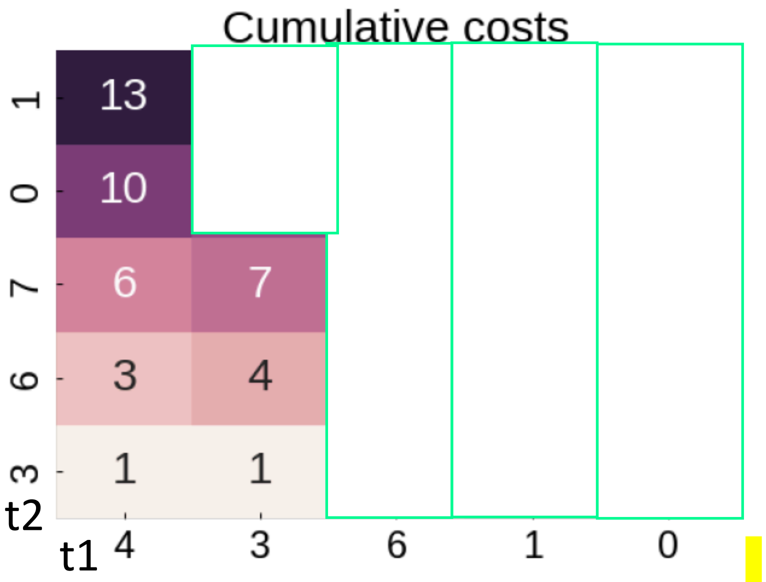
$$\gamma(i,j) = d(q_i,c_j) + \min\{\ \gamma(i-1,j-1),\ \gamma(i-1,j\ ),\ \gamma(i,j-1)\ \}$$

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)





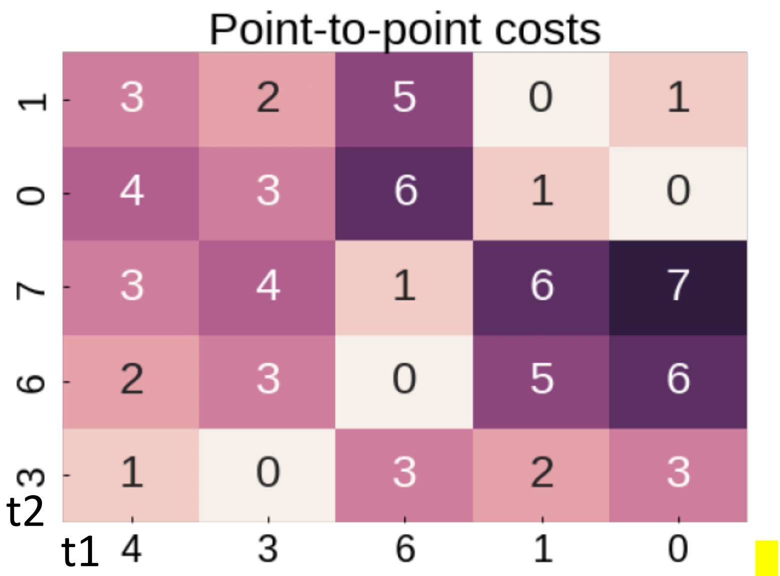
Result: 4

$$\gamma(i,j) = d(q_i,c_j) + \min\{ \gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1) \}$$

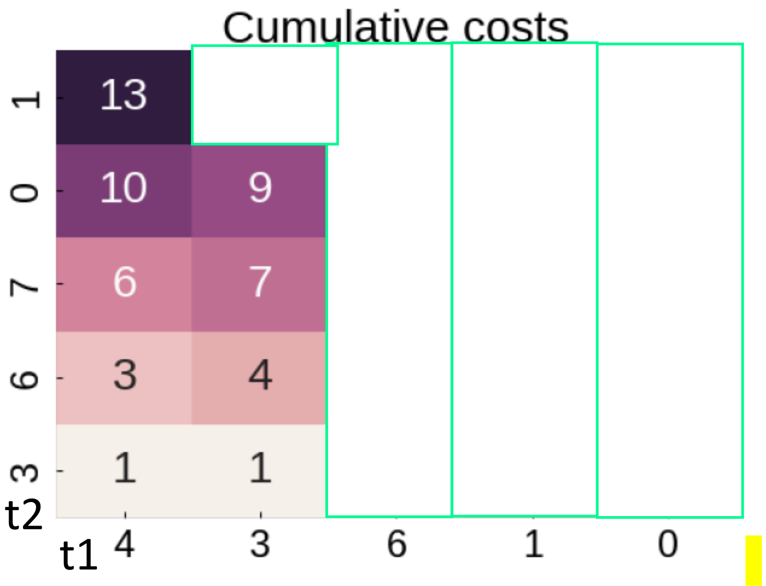
DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)



Result: 4

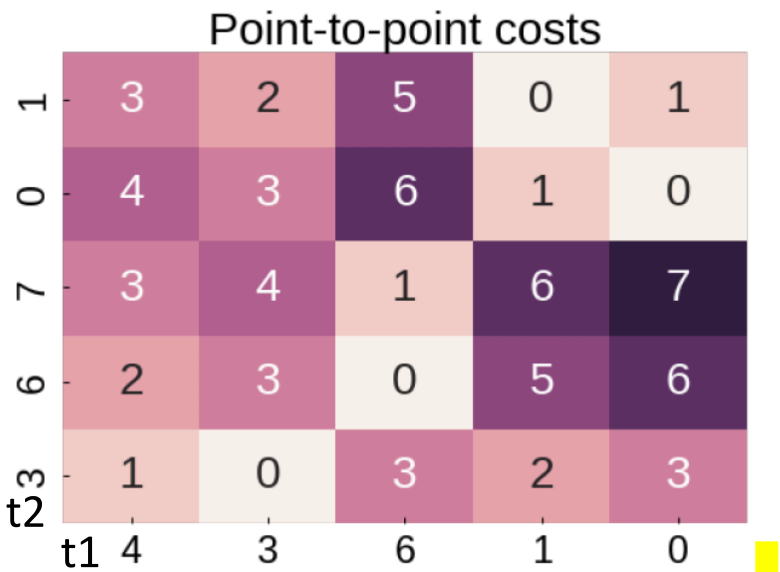


$$\gamma(i,j) = d(q_i,c_j) + \min\{ \gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1) \}$$

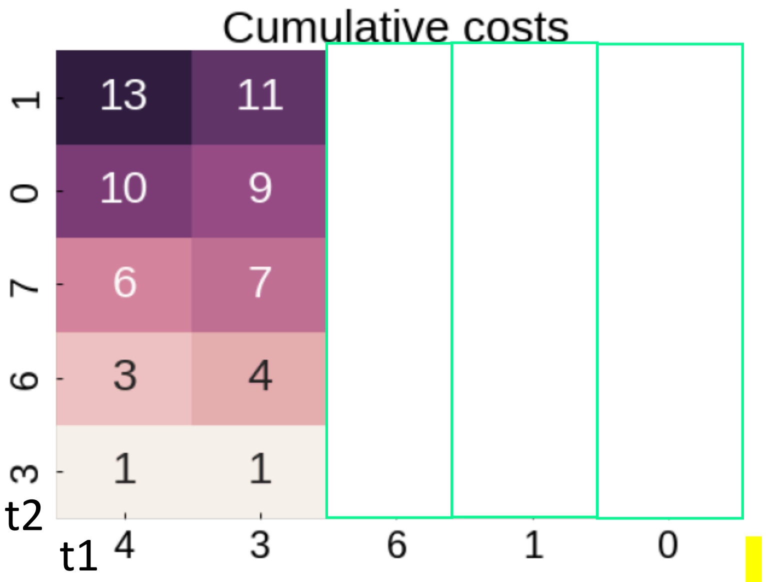
t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

DTW – Exercise 1 - Solution

- A)



Result: 4



$$\gamma(i,j) = d(q_i,c_j) + \min\{ \gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1) \}$$

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)

Point-to-point costs

t2 \ t1	4	3	6	1	0
1	3	2	5	0	1
0	4	3	6	1	0
7	3	4	1	6	7
6	2	3	0	5	6
3	1	0	3	2	3

Cumulative costs

t2 \ t1	4	3	6	1	0
1	13	11	13	3	4
0	10	9	8	3	3
7	6	7	2	7	13
6	3	4	1	6	12
3	1	1	4	6	9

Result: 4

$$\gamma(i,j) = d(q_i,c_j) + \min\{\gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1)\}$$

DTW – Exercise 1 - Solution

t1	< 4, 3, 6, 1, 0 >
t2	< 3, 6, 7, 0, 1 >

- A)

Result: 4

- B) No. Because the DTW optimal path remains inside the band of size r=1
- C) Yes. The optimal path in one direction is the same in the opposite direction. Though, the cumulative costs matrix might look different.

DTW – Exercise 2

- Given the following time series:

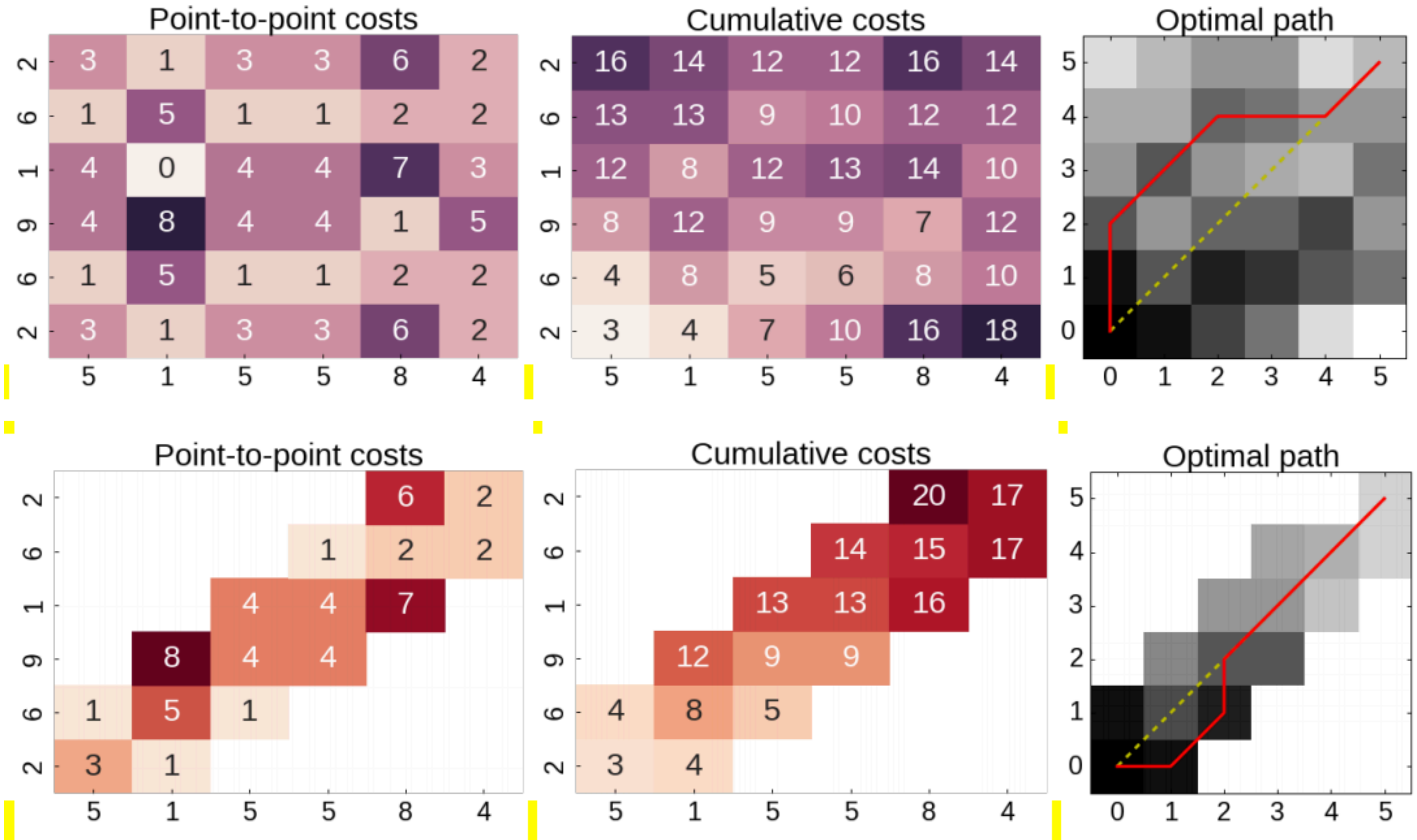
t	=	< 2, 6, 9, 1, 6, 2 >
q	=	< 5, 1, 5, 5, 8, 4 >

compute

- (i) their Manhattan and Euclidean distance,
- (ii) their DTW, and (iii) their DTW with Sakoe-Chiba band of size r=1 (i.e. all cells at distance <= 1 from the diagonal are allowed).
- For points (ii) and (iii) show the cost matrix and the optimal path found.

DTW – Exercise 2 - Solution

- Euclidean = sqrt(74) = 8.6, Manhattan = 20
- DTW = 14
- DTW r=1 = 17



DTW – Exercise 3

- Given the following time series:

ID	Time series
W	< 6, 11, 13, 15 >
X	< 10, 7, 7, 12, 14, 17 >
Y	< 9, 11, 14, 13, 20 >

- Compute the distances among all pairs of time series adopting a Dynamic Time Warping distance, and computing the distances between single points as d(x,y) = | x – y |. For each pair of time series compared also show the matrix used to compute the final result.

DTW – Exercise 3 - Solution

ID	Time series
W	< 6, 11, 13, 15 >
X	< 10, 7, 7, 12, 14, 17 >
Y	< 9, 11, 14, 13, 20 >

W – X

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	(4) 4	(1) 5	(1) 6	(6) 12	(8) 20	(11) 31
[2,]	(1) 5	(4) 8	(4) 9	(1) 7	(3) 10	(6) 16
[3,]	(3) 8	(5) 11	(5) 14	(1) 8	(1) 8	(4) 12
[4,]	(5) 13	(8) 16	(8) 19	(3) 11	(4) 9	(2) 10

W – Y

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	(3) 3	(5) 8	(8) 16	(7) 23	(14) 37
[2,]	(2) 5	(0) 3	(3) 6	(2) 8	(9) 17
[3,]	(5) 9	(2) 5	(1) 4	(0) 4	(7) 11
[4,]	(6) 15	(4) 9	(1) 5	(2) 6	(5) 9

X – Y

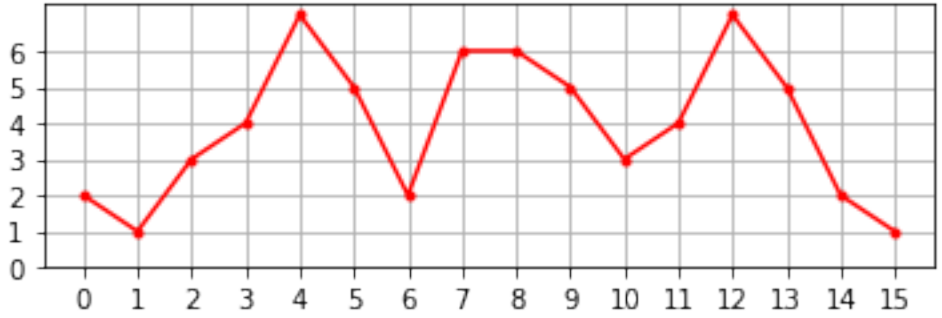
	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	(1) 1	(1) 2	(4) 6	(3) 9	(10) 19
[2,]	(2) 3	(4) 5	(7) 9	(6) 12	(13) 22
[3,]	(2) 5	(4) 7	(7) 12	(6) 15	(13) 25
[4,]	(3) 8	(1) 6	(2) 8	(1) 9	(8) 17
[5,]	(5) 13	(3) 9	(0) 6	(1) 7	(6) 13
[6,]	(8) 21	(6) 15	(3) 9	(4) 10	(3) 10

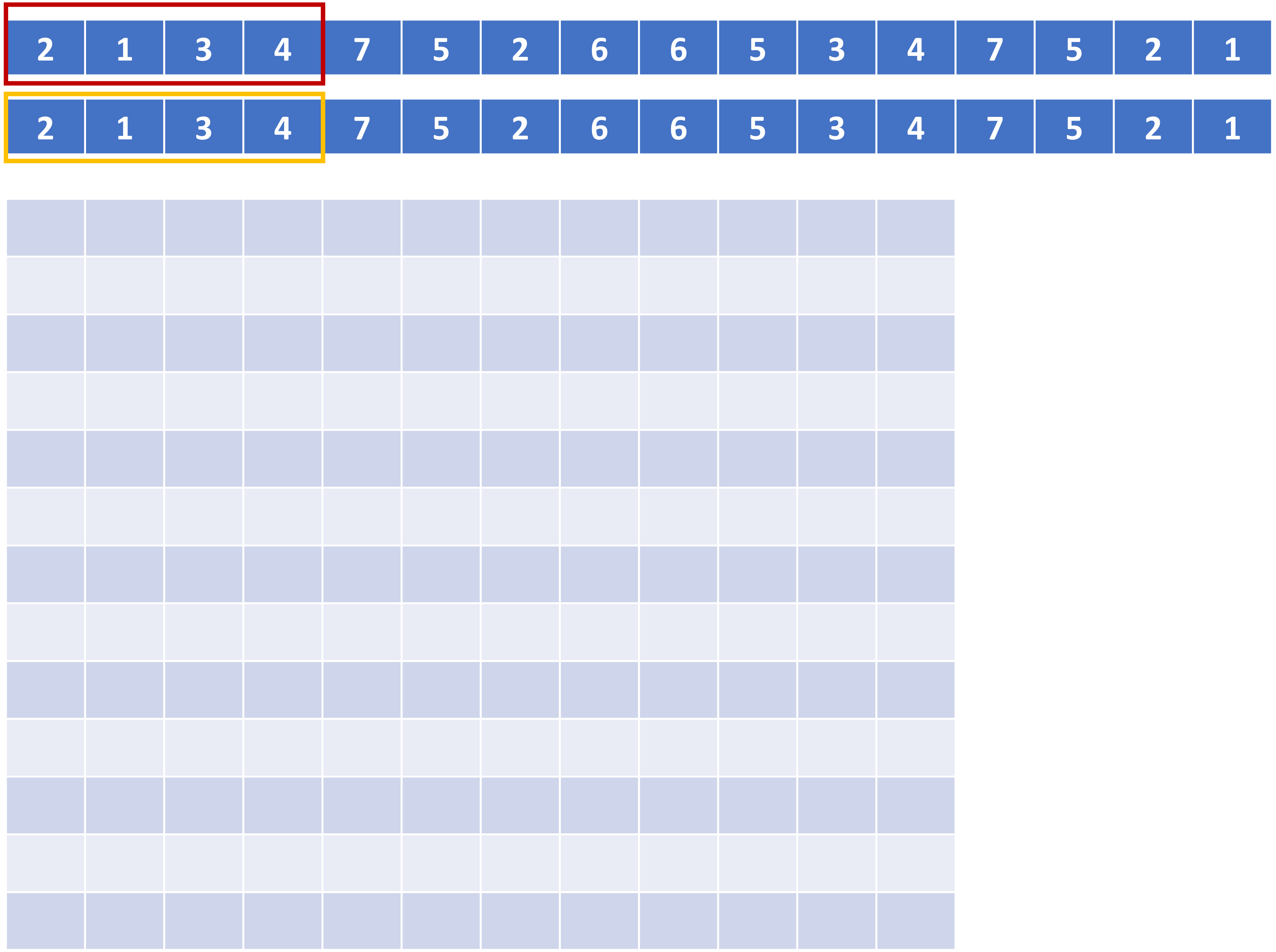
Matrix Profile

Matrix Profile

Given the TS x = <2,1,3,4,7,5,2,6,6,5,3,4,7,5,2,1> and the

1. Build the Matrix Profile for x with m=4 using the Manahttan distance as distance function between subsequences.
2. Draw the Matrix Profile
3. Identify the motifs with distance equals 0 and length equals to m
4. Which is a correct value for m that would have retrieved less motifs with distance equals to 0?





m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7											

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9										

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12									

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12	12								

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12	12	9	9	12	8	9	9	12	13

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12	12	9	9	12	8	9	9	12	13

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12	12	9	9	12	8	9	9	12	13
7											

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12	12	9	9	12	8	9	9	12	13
7	8										

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12	12	9	9	12	8	9	9	12	13
7	8	13									

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12	12	9	9	12	8	9	9	12	13
7	8	13	11								

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1
2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12	12	9	9	12	8	9	9	12	13
7	8	13	11	8	8	13	11	4	8	13	16

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12	12	9	9	12	8	9	9	12	13
7	8	13	11	8	8	13	11	**4**	8	13	16

m = 4

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12	12	9	9	12	8	9	9	12	13
7	8	13	11	8	8	13	11	**4**	8	13	16
9											

m = 4

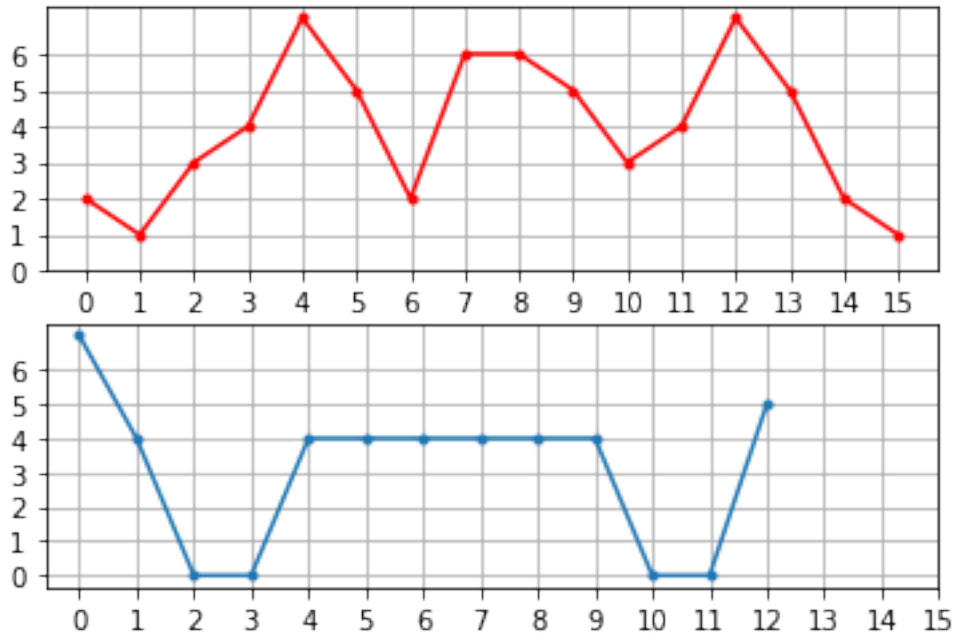
2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1
2	1	3	4	7	5	2	6	6	5	3	4	7	5	2	1

7	9	12	12	9	9	12	8	9	9	12	13
7	8	13	11	8	8	13	11	**4**	8	13	16
9	8	9	11	6	4	9	9	8	**0**		

m = 4

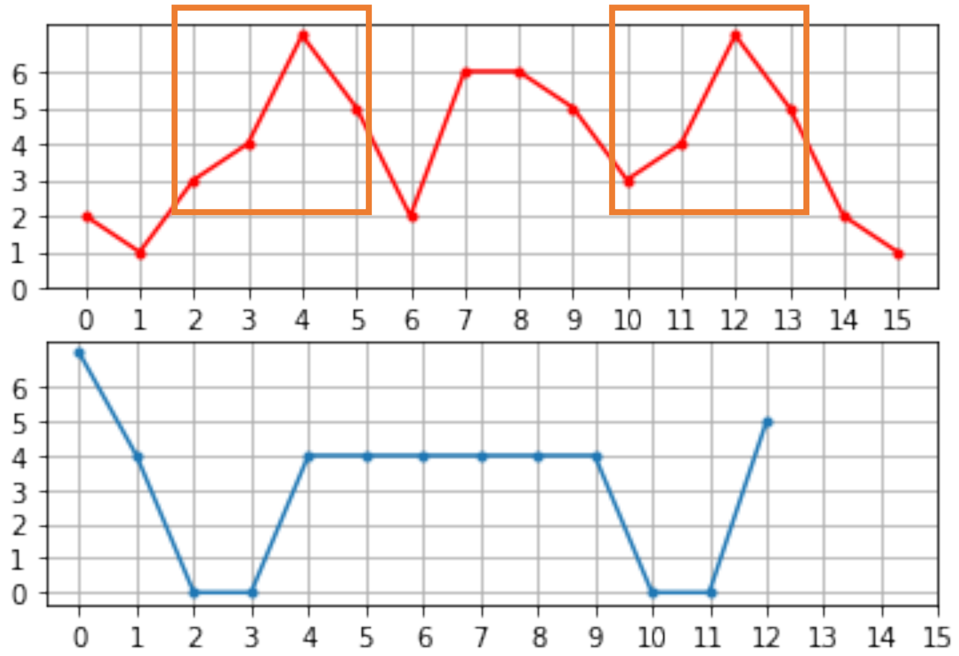
Matrix Profile

- x = <2, 1, 3, 4, 7, 5, 2, 6, 6, 5, 3, 4, 7, 5, 2, 1>
- mp = < 7, 4, 0, 0, 4, 4, 4, 4, 4, 4, 0, 0, 5 >



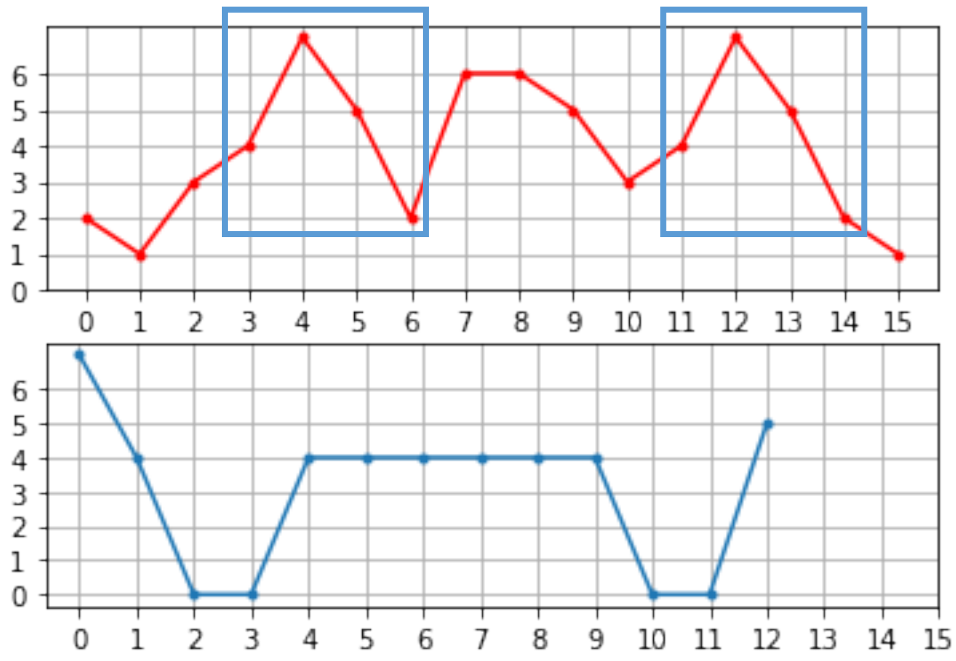
Matrix Profile

- x = <2, 1, 3, 4, 7, 5, 2, 6, 6, 5, 3, 4, 7, 5, 2, 1>
- mp = < 7, 4, 0, 0, 4, 4, 4, 4, 4, 4, 0, 0, 5 >



Matrix Profile

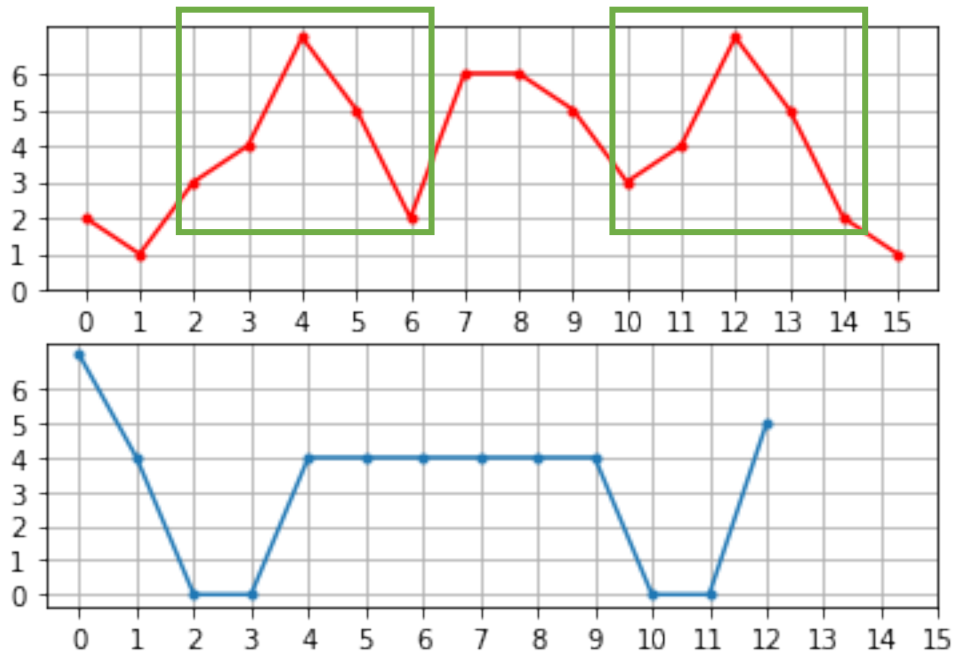
- x = <2, 1, 3, 4, 7, 5, 2, 6, 6, 5, 3, 4, 7, 5, 2, 1>
- mp = < 7, 4, 0, 0, 4, 4, 4, 4, 4, 4, 0, 0, 5 >



Matrix Profile

- x = <2, 1, 3, 4, 7, 5, 2, 6, 6, 5, 3, 4, 7, 5, 2, 1>
- mp = < 7, 4, 0, 0, 4, 4, 4, 4, 4, 4, 0, 0, 5 >

with m = 5



Matrix Profile

Given the TS x = <5,5,3,5,5,1> and the

1. Build the Matrix Profile for x with m=2 using the Manahttan distance as distance function between subsequences.
2. Draw the Matrix Profile
3. Identify the motifs with distance equals 0 and length equals to m