Machine Learning Data Analytics

FAU

Data Mining

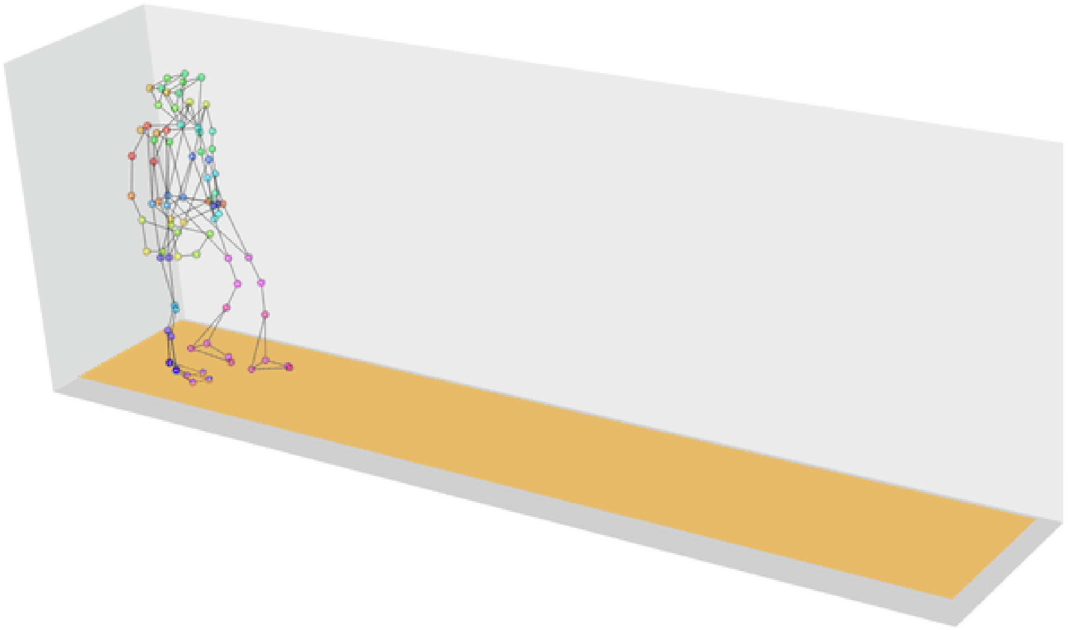
Dynamic Time Warping

Richard Dirauf, M.Sc.
Machine Learning and Data Analytics (MaD) Lab
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)
MLTS Exercise, 12.01.2024

- Data characteristics, reduction, transformation
- Frequency analysis
    - Spectral analysis
    - Continuous Wavelet Transform
- **Dynamic Time Warping**
- Similarity join, Matrix profile, Signature method

In many application, there is the need to analyze multiple time series at the same time to find similarities between time series.

- E.g., speech recognition, signature recognition, similarity in walking, …

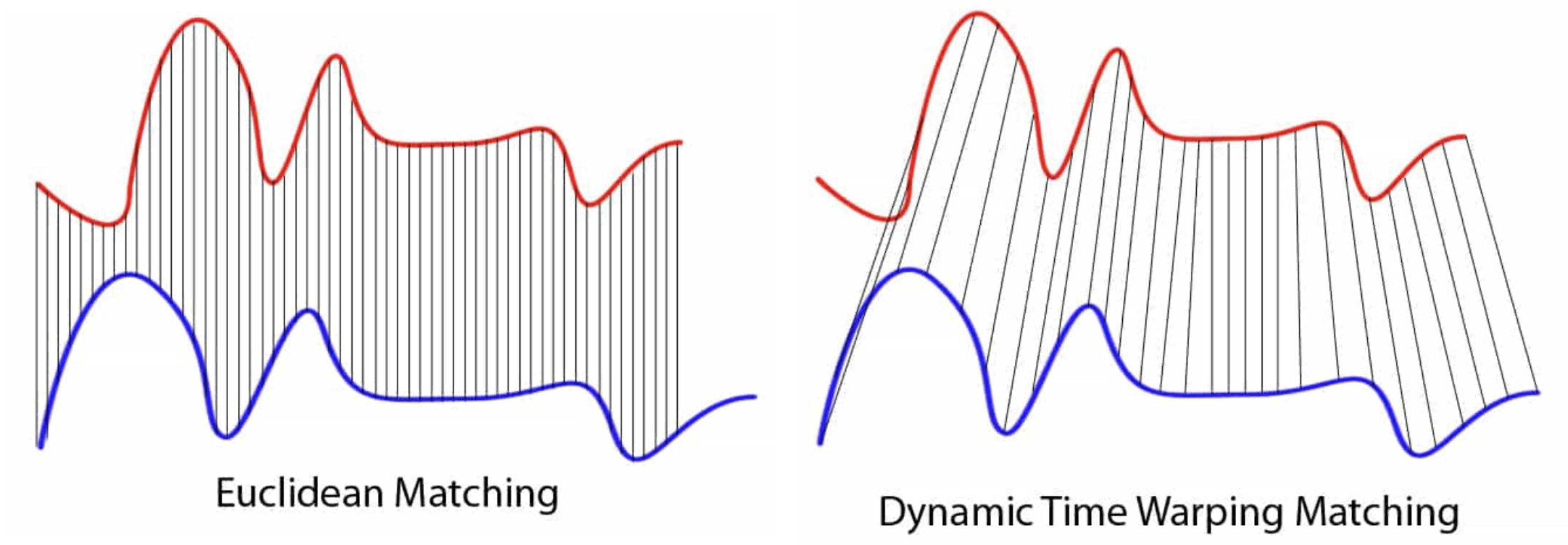


https://en.wikipedia.org/wiki/Motion_capture

Dynamic time warping

MLTS Exercise 06

Euclidean distance does not work for time series that are not perfectly synchronized.

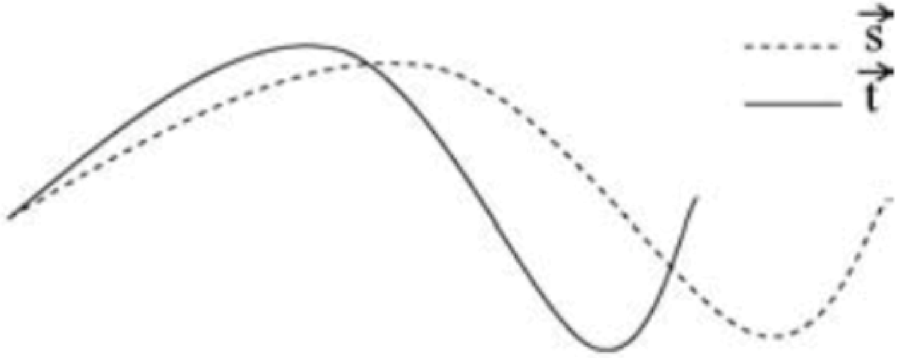


Dynamic time warping

MLTS Exercise 06

Dynamic time warping (DTW) is an algorithm to measure similarity between two time-series that may **vary in speed and time**.

DTW determines the optimal global alignment between two time series.



Use cases

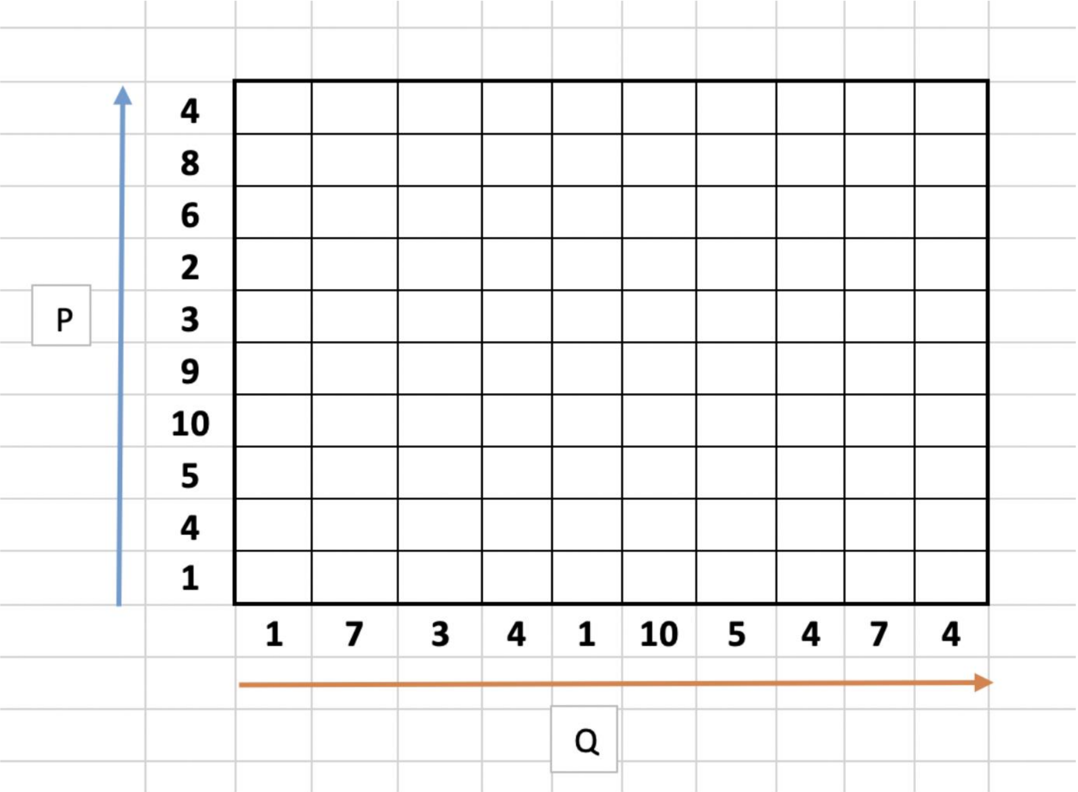
- *Financial markets* – comparing stock trading data over similar time frames. For example, comparing monthly trading data for February (28 days) and March (31 days).
- *Wearable fitness trackers* – more accurately calculating a walker’s speed and the number of steps, even if their speed varied over time.
- *Route calculation* – calculating more accurate information about a driver’s ETA, if we know something about their driving habits (for example, they drive quickly on straightaways but take more time than average to make left turns).

Dynamic time warping: Algorithm

MLTS Exercise 06

Given two time series p_t and q_t, we can compute the DTW distance as follows:

1. Initialize the distance matrix M



Dynamic time warping: Algorithm

MLTS Exercise 06

Given two time series p_t and q_t ,we can compute the DTW distance as follows:

1. Initialize the distance matrix M
2. Fill M from the bottom left corner, according to the formula:

$$M(i,j) = dist\left(x_i, y_j\right) + \min(M(i-1, j-1), M(i, j-1), M(i-1, j))$$

P

4	42	24	-	-	-	-	-	-	-	-
8	39	21	-	-	-	-	-	-	-	-
6	32	20	-	-	-	-	-	-	-	-
2	27	19	-	-	-	-	-	-	-	-
3	26	14	-	-	-	-	-	-	-	-
9	24	10	-	-	-	-	-	-	-	-
10	16	8	12	**11**	-	-	-	-	-	-
5	7	5	5	5	-	-	-	-	-	-
4	**3**	3	4	4	7	13	14	14	17	17
1	0	6	**8**	11	11	20	24	27	33	36
	1	**7**	**3**	**4**	**1**	**10**	**5**	**4**	**7**	**4**

Q

Dynamic time warping: Algorithm

MLTS Exercise 06

Given two time series p_t and q_t ,we can compute the DTW distance as follows:

1. Initialize the distance matrix M
2. Fill M from the bottom left corner, according to the formula:
   $M(i,j) = dist\left(x_i, y_j\right) + \min(M(i-1, j-1),$
   $M(i, j-1), M(i-1, j))$
3. Identify the warping path d, starting from the top right corner

P										
4	42	24	20	17	20	25	19	18	18	**15**
8	39	21	19	17	20	19	18	19	**15**	18
6	32	20	14	13	17	18	15	15	**14**	16
2	27	19	11	12	12	20	14	**13**	17	18
3	26	14	10	11	13	16	**11**	12	16	17
9	24	10	14	16	19	**9**	12	17	18	21
10	16	8	12	11	14	**8**	13	18	16	21
5	7	5	5	5	**8**	12	12	13	15	16
4	3	**3**	**4**	**4**	7	13	14	14	17	17
1	**0**	6	8	11	11	20	24	27	33	36
	1	**7**	**3**	**4**	**1**	**10**	**5**	**4**	**7**	**4**
					Q					

Dynamic time warping: Algorithm

MLTS Exercise 06

Given two time series p_t and q_t ,we can compute the DTW distance as follows:

1. Initialize the distance matrix M
2. Fill M from the bottom left corner, according to the formula:
   $M(i,j) = dist(x_i, y_j) + \min(M(i-1, j-1),$
   $M(i, j-1), M(i-1, j))$
3. Identify the warping path d, starting from the top right corner
4. Calculate the final distance D as

$$D = \sum_{(i,j) \in d} dist(x_i, y_j)$$

P										
4	42	24	20	17	20	25	19	18	18	**15**
8	39	21	19	17	20	19	18	19	**15**	18
6	32	20	14	13	17	28	15	15	**14**	16
2	27	19	11	12	12	20	14	**13**	17	18
3	26	14	10	11	13	16	**11**	12	16	17
9	24	10	14	16	19	**9**	12	17	18	21
10	16	8	12	11	14	**8**	13	18	16	21
5	7	5	5	5	**8**	12	12	13	15	16
4	3	**3**	**4**	**4**	7	13	14	14	17	17
1	**0**	6	8	11	11	20	24	27	33	36
	1	**7**	**3**	**4**	**1**	**10**	**5**	**4**	**7**	**4**

Q

Dynamic time warping: Pros and Cons

MLTS Exercise 06

Pros:

- Exploit a non-linear distortion (in time) to find non-trivial similarity

Cons:

- **High computational cost.** Alternatives for computing the alignment path more efficiently have been presented.
- It needs the preparation of reliable reference templates for the set of words to be recognized.

Practice Question

MLTS Exercise 08

Dynamic Time Warping Task

- Given the two time-series

  P = [1, 9, 5, 7] and Q = [3, 8, 2, 4]

- Fill the distance matrix M
- Identify the warping path d
- Calculate the final distance D along the warping path

Any questions?