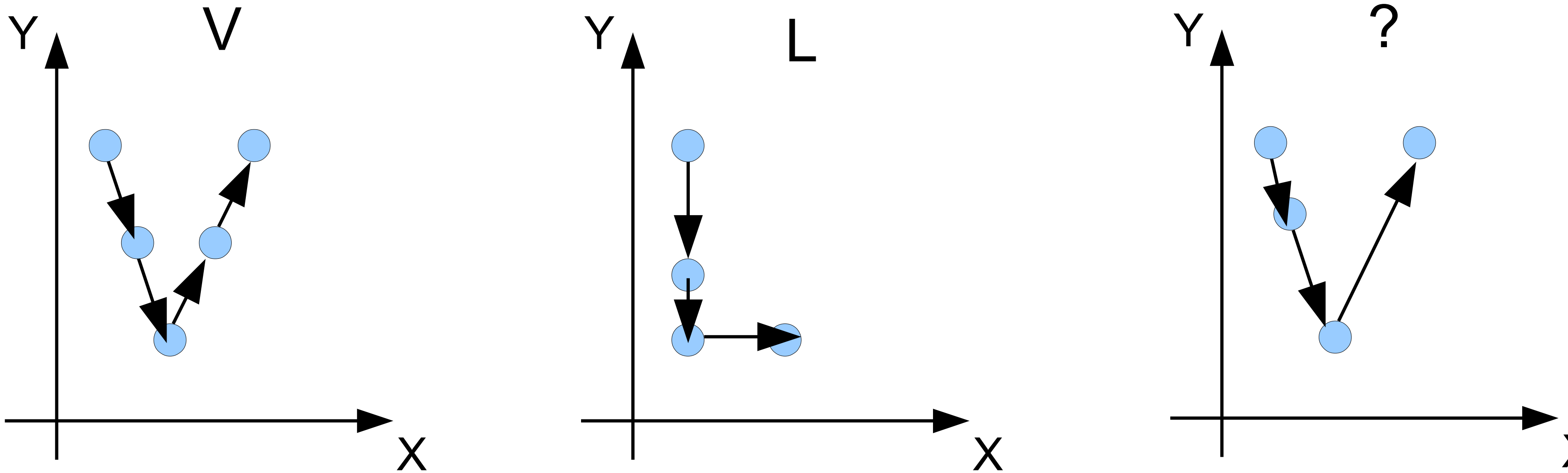
Mining and Processing Biomedical Data

Dr. rer. nat. Krisztian Buza
adiunkt naukowy
Faculty of Mathematics, Informatics and Mechanics
University of Warsaw, Poland
chrisbuza@yahoo.com

Multivariate DTW

Example: handwriting recognition



Time series (deltaX, deltaY):

(1,-2), (1, -2), (1, 2), (1, 2)   (0,-3), (0, -1), (3, 0)   (0.5,-1), (1.5, -3), (2, 4)

DTW on multivariate time series

	(0.5, -1),	(1.5, -3),	(2, 4)
(1, -2)	**1.118**		
(1, -2)			
(1, 2)			
(1, 2)			

$$\sqrt{(1 - 0.5)^2 + ((-2) - (-1))^2}$$

DTW on multivariate time series

	$\boxed{(0.5, -1),}$	(1.5, -3),	(2, 4)
(1, -2)	**1.118**		
$\boxed{(1, -2)}$	**2.236**		
(1, 2)			
(1, 2)			

$$1.118 + \sqrt{(1-0.5)^2 + ((-2) - (-1))^2}$$

DTW on multivariate time series

	$\boxed{(0.5, -1)}$,	(1.5, -3),	(2, 4)
(1, -2)	**1.118**		
(1, -2)	**2.236**		
$\boxed{(1, 2)}$	**5.277**		
(1, 2)			

$$2.236 + \sqrt{(1 - 0.5)^2 + (\,2 - \,(-1)\,)^2}$$

DTW on multivariate time series

	(0.5, -1),	(1.5, -3),	(2, 4)
(1, -2)	**1.118**		
(1, -2)	**2.236**		
(1, 2)	**5.277**		
(1, 2)	**8.318**		

$$5.277 + \sqrt{(1 - 0.5)^2 + (2 - (-1))^2}$$

DTW on multivariate time series

	(0.5, -1),	(1.5, -3),	(2, 4)
(1, -2)	**1.118**	**2.236**	
(1, -2)	**2.236**		
(1, 2)	**5.277**		
(1, 2)	**8.318**		

$$1.118 + \sqrt{(1 - 1.5)^2 + ((-2) - (-3))^2}$$

DTW on multivariate time series

	(0.5, -1),	(1.5, -3),	(2, 4)
(1, -2)	**1.118**	**2.236**	
(1, -2)	**2.236**	**2.236**	
(1, 2)	**5.277**		
(1, 2)	**8.318**		

$$1.118 + \sqrt{(1 - 1.5)^2 + ((-2) - (-3))^2}$$

Min {1.118, 2.236, 2.236}

DTW on multivariate time series

	(0.5, -1),	(1.5, -3),	(2, 4)
(1, -2)	**1.118**	**2.236**	
(1, -2)	**2.236**	**2.236**	
(1, 2)	**5.277**	**7.261**	
(1, 2)	**8.318**		

$$2.236 + \sqrt{(1 - 1.5)^2 + (\ 2 - \ (-3)\ )^2}$$

Min {5.277, 2.236, 2.236}

DTW on multivariate time series

	(0.5, -1),	$\boxed{\text{(1.5, -3),}}$	(2, 4)
(1, -2)	**1.118**	**2.236**	
(1, -2)	**2.236**	**2.236**	
(1, 2)	**5.277**	**7.261**	
$\boxed{\text{(1, 2)}}$	**8.318**	**12.286**	

$$5.277 + \sqrt{(1 - 1.5)^2 + (\,2 - \;(-3)\,)^2}$$

Min {5.277, 7.261, 8.318}

DTW on multivariate time series

	(0.5, -1),	(1.5, -3),	(2, 4)
(1, -2)	**1.118**	**2.236**	**8.319**
(1, -2)	**2.236**	**2.236**	**8.319**
(1, 2)	**5.277**	**7.261**	**4.472**
(1, 2)	**8.318**	**12.286**	**6.708**

- What if the time-series have more than two channels (ECG, EEG)?
- Euclidean distance vs. cosine distance

Speeding-up nearest neighbor classification of time-series

Speeding-up nearest neighbor classifiation of time-series

- For accurate classification, in many cases, large training dataset is required → the classifier may become slow

Speeding-up nearest neighbor classifiation of time-series

- For accurate classification, in many cases, large training dataset is required → the classifier may become slow
- Speed-up techniques
    - Efficient computation of the similarity / distance of time series

Speeding-up nearest neighbor classifiation of time-series

- For accurate classification, in many cases, large training dataset is required → the classifier may become slow
- Speed-up techniques
  - Efficient computation of the similarity / distance of time series
  - Avoiding the computation of all the distances (lower bounding, early stopping of DTW-computation)

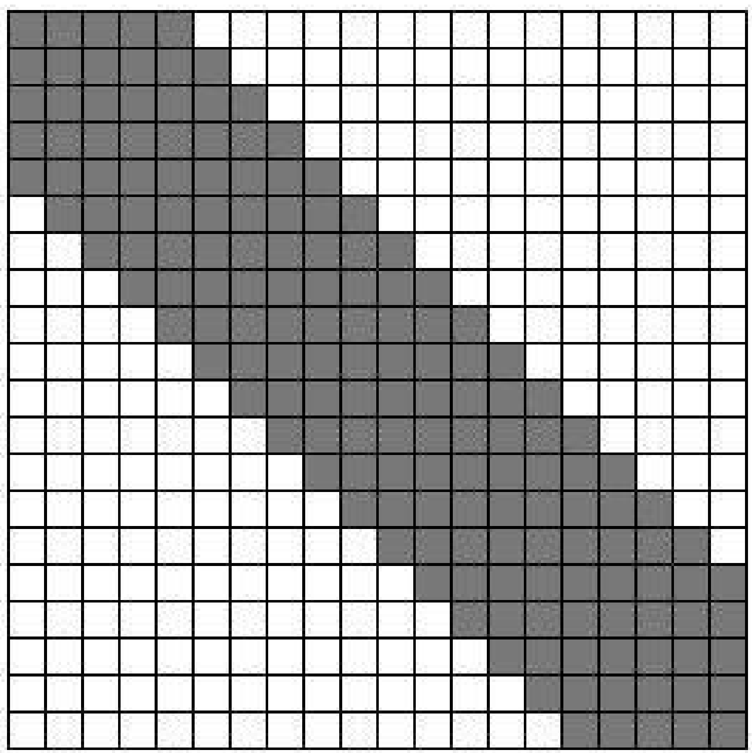
Speeding-up nearest neighbor classifiation of time-series

- For accurate classification, in many cases, large training dataset is required → the classifier may become slow
- Speed-up techniques
  - Efficient computation of the similarity / distance of time series
  - Avoiding the computation of all the distances (lower bounding, early stopping of DTW-computation)
  - Preprocessing techniques (esp. reduction of the length by the aggregation of consequtive values)

Speeding-up nearest neighbor classifiation of time-series

- For accurate classification, in many cases, large training dataset is required → the classifier may become slow
- Speed-up techniques
  - Efficient computation of the similarity / distance of time series
  - Avoiding the computation of all the distances (lower bounding, early stopping of DTW-computation)
  - Preprocessing techniques (esp. reduction of the length by the aggregation of consequtive values)
  - Numerosity reduction / instance selection

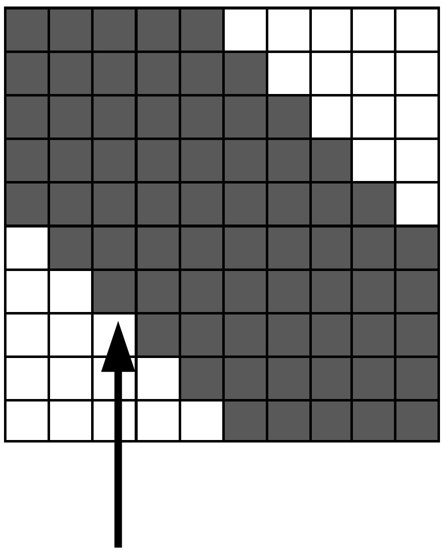
Constrained DTW



- Calculate only the marked entries of the DTW-matrix, i.e., the ones that are „close“ to the diagonal of the matrix
- „Extreme“ variant of this idea: Lucky Time Warping (Spiegel, 2014)

Spiegel, Stephan, Brijnesh-Johannes Jain, and Sahin Albayrak. "Fast Time Series Classification under Lucky Time Warping Distance." SAC 2014.

Early Stop



This column was just calculated. If all the entries in this column are larger than d, we do not need to calculate the rest of the matrix.

- We want to determine the nearest neighbors of the time series t^*
- We are in an intermediate step, i.e., we already calculated the distance between t^* and some of the time series of the training data → we know that the distance between t^* and some other time series is d
- We are calulating the distance between t^* and the currently considered time series.
- If the DTW matrix has only entries being greater than d in the column that was calculated last → stop the calculations (the currently considered time series can not be the nearest neighbor anyway).

Nearest Neighbor Classification with Lower Bounding

*t** – Time series to be classified
*d** – distance of the currently found closest time series

$d^* \leftarrow$ infinity
for each time series t of the training data {
    $d \leftarrow$ estimate_distance(t^*, t)
    (d is a lower bound, i.e., the estimation is done in a
      way that the true distance is greater than or equal to d)
    if ($d > d^*$) **continue**;
    $d1 \leftarrow$ DTW(t^*, t)
    if ($d1 < d^*$) {
        $d^* \leftarrow d1$; $nearest\_neighbor \leftarrow t$;
    }
}

See also:
Rath, Toni M., and R. Manmatha. "Lower-bounding of dynamic time warping distances for multivariate time series." (2003).