Testing the correlation of time series using dynamic time warping

N. Pandria
Medical School, Aristotle University of Thessaloniki

D. Kugiumtzis
Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki

27o Panhellenic Conference of Statistics, Thessaloniki
The problem

Are any of the three time series “similar” to each other?

Outline
- Align the time series (DTW)
- Significance correlation test
- Simulations
- Application on TMS-EEG

\[
\begin{align*}
  r(X,Y) &= -0.25 \\
  r(X,Z) &= -0.12 \\
  r(Y,Z) &= -0.38
\end{align*}
\]
Dynamic Time Warping Algorithm (DTW)

1. Assume two time series:
   \[ Q = q_1, q_2, \ldots, q_n \]
   \[ C = c_1, c_2, \ldots, c_m \]

2. Generate a \( n \times m \) distance matrix using a distance measure, e.g. Euclidean:
   \[ d(q_i, c_j) = (q_i - c_j)^2 \]

3. Define the warping path: \( W = w_1, w_2, \ldots, w_k, \ldots, w_K \) where \( \max(n, m) \leq K < n + m - 1 \) satisfying the following conditions:
   - **Boundary conditions**: \( w_1 = (1, 1) \) and \( w_K = (n, m) \)
   - **Continuity**: if \( w_k = (a, b) \) and \( w_{k-1} = (a', b') \), then \( a - a' \leq 1 \) and \( b - b' \leq 1 \)
   - **Monotonicity**: if \( w_k = (a, b) \) and \( w_{k-1} = (a', b') \), then \( a - a' \geq 0 \) and \( b - b' \geq 0 \)

The optimal warping path can be found through dynamic programming evaluating the following recurrent equation:

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

The DTW distance measure is:

\[ \text{DTW}(Q, C) = \gamma(n, m) \]
DTW in practice

- Compute $\gamma(1,1) = d(q_1, c_1)$.

- Evaluate the 1st row of the warping matrix
  
  \[ \gamma(i, 1) = \gamma(i-1, 1) + d(q_i, c_1). \]

- Compute the 1st column
  
  \[ \gamma(1, j) = \gamma(1, j-1) + d(q_1, c_j). \]

- Move to the 2nd row
  
  \[ \gamma(i, 2) = \min\{\gamma(i, 1), \gamma(i-1, 1), \gamma(i-1, 2)\} + d(q_i, c_2) \]

- Moving from left to right we evaluate the recurrence:
  
  \[ \gamma(i, j) = \min\{\gamma(i, j-1), \gamma(i-1, j-1), \gamma(i-1, j)\} + d(q_i, c_j). \]

- Finally, we move backwards from $\gamma(n,m)$ to $\gamma(1,1)$ following the red arrows of minimum cost.

- The two aligned time series are formed from the time indices along the optimal warping path.

- The path length $N$ may be $N > \max(n,m)$
Cross-correlation and DTW

Two time series \( X \) and \( Y \) may not align or may even not have the same length.

How can their similarity be assessed?

Cross-correlation

\[
c_{X,Y}(\tau) = \frac{1}{n} \sum_{t=1}^{n-\tau} (x_t y_{t+\tau} - \bar{x} \bar{y})
\]

\[
r_{X,Y}(\tau) = \frac{c_{X,Y}(\tau)}{s_X s_Y}
\]

and particularly

\[
r_{X,Y}(0) \equiv r(X,Y)
\]

… but \( X \) and \( Y \) may not have the same length?

Solution: \( (X,Y) \xrightarrow{\text{DTW}} (X^d,Y^d) \xrightarrow{} r(X^d,Y^d) \)

Significance of \( r(X^d,Y^d) \)?

The alignment of DTW is expected to increase the correlation, \( r(X^d,Y^d) > r(X,Y) \)

Parametric significance test is not valid

Resampling (randomization) significance test, \( H_0: \rho(X^d,Y^d)=0 \)
Cross-correlation and DTW

DTW on $X$ and $Y$

DTW on $X$ and $Z$

$r(X,Y) = -0.25$

$r(X,Z) = -0.12$

$r(Y,Z) = -0.38$

$r(X^d,Y^d) = 0.80$

very large $r$

$r(X^d,Y^d) = 0.61$
Randomization significance test for cross-correlation

Two time series $X^d$ and $Y^d$ of length $N$.

**The problem:**
Generate randomized copies $(X^{*d}, Y^{*d})$ of $(X^d, Y^d)$ consistent to $H_0$: $\rho(X^d, Y^d)=0$

$$\rho(X^{*d}, Y^{*d})=0 \text{ and } X^{*d} \text{ is “similar” to } X^d \text{ and } Y^{*d} \text{ is “similar” to } Y^d$$

$X^{*d}$ has the same **marginal distribution** as $X^d$

Random permutation (RP)

$X^{*d}$ has the same **marginal distribution** and the same **linear structure** as $X^d$

Iterated amplitude adjusted Fourier transform (IAAFT) [2] preserves the original power spectrum

Statically transformed autoregressive process (STAP) [3] preserves the original autocorrelation
Randomization significance test for cross-correlation

The randomization test for $H_0: \rho(X,Y)=0$ of two time series $X$ and $Y$ of length $N$.

1. Generate $M$ randomized copies of $X$ and $Y$ (using RP, IAAFT or STAP)

2. Compute the cross-correlation (for lag zero):
   - on the original pair $r_0=r(X,Y)$
   - and on each of the $M$ randomized pairs $(X^*_i, Y^*_i), i=1,\ldots,M$
     
     $r_i=r(X^*_i, Y^*_i), i=1,\ldots,M$

3. Reject $H_0$ if $r_0$ is at the tails of the sample distribution of $\rho(X,Y)$ formed by $r_1, r_2,\ldots, r_M$ [we use rank ordering]

For significance level $\alpha$, if the rank of $r_0$ in the ordered list of $r_0, r_1, r_2,\ldots, r_M$ is $<(M+1)\alpha/2$ or $>(M+1)(1-\alpha/2)$ then $H_0$ is rejected.
Simulations

Setup

1. Generate two time series $X'$ and $Y'$ (with / without autocorrelation and cross-correlation) $(X',Y') \sim \text{VAR}(1)$

2. Make single gaps in the one time series ($X'$) at a given percentage of total length, denoted $X$, while the other is intact ($Y\equiv Y'$).

3. $r(X',Y') = r^0 \leftarrow \{X',Y'\} \rightarrow \{X,Y\} \rightarrow \text{DTW} \rightarrow \{X^d,Y^d\} \rightarrow r(X^d,Y^d) = r^{d,0}$

4. Generate $M=1000$ surrogates of $X$ using each of RP, IAAFT and STAP ($Y$ is intact), $X^*i$ for $i=1,\ldots,M$

5. $r(X^*i,Y) = r^*i \leftarrow \{X^*i,Y\} \rightarrow \text{DTW} \rightarrow \{X^{*id},Y^d\} \rightarrow r(X^{*id},Y^d) \equiv r^{*d,i}$ for $i=1,\ldots,M$

6. From the placement of $r^0$ in the ordered list $\{r^0, r^*1, r^*2, \ldots, r^*M\}$ make the test decision at the significance level $\alpha=0.05$ $H_0: \rho(X,Y) = 0$

7. From the placement of $r^{d,0}$ in the ordered list $\{r^{d,0}, r^{*d,1}, r^{*d,2}, \ldots, r^{*d,M}\}$ make the test decision at the significance level $\alpha=0.05$ $H_0: \rho(X^d,Y^d) = 0$
Example

1. Generate single gaps in $X'$ (20% x length($X'$)) while $Y'$ remains stable.

2. Scatter plot and cross-correlation of ($X'$, $Y'$)

$X$ and $Y$ generated by a VAR(1) model

Scatter plot and cross-correlation of ($X'$, $Y'$) before DTW CA = 0.1786
Example (continues)

Scatter plot and cross-correlation of \((X^d, Y^d)\)

3. DTW
Example (continues)

Generation of surrogates for $X$

RP

IAAFT

STAP

Cross-correlation of $(X^{*i}, Y)$ BEFORE DTW

Cross-correlation of $(X^{*id}, Y^d)$ AFTER DTW

Significance Test
Example (continues)

Before DTW

After DTW

1000 surrogates, p-value=0.971365 before DTW

1000 surrogates, p-value=0.843537 after DTW

Legend:
- 25%
- 75%
- orig
- surr
Simulation Results

- Independent and uncorrelated time series

Both cross-correlation and autocorrelation are zero

All tests perform properly: no rejection

Before DTW

After DTW
• Dependent and uncorrelated time series

Significant autocorrelation but zero cross-correlation

Before DTW

After DTW

RP fails

weak autocorrelation

strong autocorrelation
• Dependent and correlated time series

Significant autocorrelation and cross-correlation

Before DTW

After DTW

All tests perform properly: rejection
Aggregate Results of surrogate generation algorithms

RP

IAAFT

STAP
Application to TMS-EEG

In collaboration with Vassilis Kimiskidis, Laboratory of Clinical Neurophysiology, AHEPA Hospital, Medical School

DTW in conjunction with significance test of cross-correlation was applied to EEG after TMS administration in order to investigate the following question:

“Is there homogeneity between different post-TMS recordings of the same individual?”

TMS: transcranial magnetic stimulation
EEG: electroencephalogram
Experimental setup

- 24 EEG recordings from patients with epilepsy
- The analysis was done on two channels (C3, CP3).
- Each recording has 15 epochs of administration of a block of 2 TMS.
- EEG segments of 1.4 sec after second TMS were extracted for analysis (sampling time is 1/1450 sec).
- The procedure of DTW, significance cross-correlation test was applied to all possible couples of EEG segments, i.e.
  - all epochs within one recording (Pintra),
  - all epochs across different recordings (Pinter)
- We test whether the postTMS EEG segments of the two recordings of one patient (P7-P8) are similar (Pintra for this patient) as compared to all other postTMS EEG segments (Pinter).
All epochs (time series) from both P7 and P8 recordings were saved in a matrix 2048 x 30.

For each pair of time series DTW applied and Pearson correlation coefficient was computed.

An upper triangular correlation matrix for each pair of warped time series was created.

The Intraepochs and Interepochs blocks were extracted as vectors.

The AUROC was measured as well as the histograms were constructed for each category (Pall-Pintra, Pall-Pinter, Pinter-Pintra).
Conclusions

- DTW can reveal similarities between time series
- The time series obtained by DTW have increased cross-correlation
- An appropriate significance test for cross-correlation is proposed using randomization (surrogate time series)
- This approach could find similarities in the response of the same person (patient with epilepsy) to the same stimulus (transcranial magnetic stimulation).

Open problems:
- Improve DTW to make a better alignment
- Can we develop a procedure that could identify patient-specific response to TMS?
Conclusions of Real-Data Application

- The distributions of Pinter and Pintra through DTW in conjunction with cross-correlation were shown to be distinct indicating the heterogeneity of the same subject’s recordings.

- In contrary, the simple Pearson correlation coefficient failed to detect the difference between the distributions of Pinter and Pintra since the auroc value is about 0.5.

- The absence of homogeneity across the recordings of the same subject might render impossible the distinction of recordings between healthy and patient individuals.


Thank you for your attention