Aligning Time Series on Incomparable Spaces

Samuel Cohen[1] · Giulia Luise[2] · Alexander Terenin[2] · Brandon Amos[3] · Marc Peter Deisenroth[1]

[1]Centre for Artificial Intelligence, University College London · [2]Imperial College London · [3]Facebook AI Research

Abstract

Dynamic time warping (DTW) is a useful method for aligning, comparing and combining time series, but it requires them to live in comparable spaces. To alleviate this, we propose Gromov dynamic time warping (GDTW), a distance between time series on potentially incomparable spaces that avoids the comparability requirement by instead considering intra-relational geometry, and automatically incorporates invariances. It draws inspiration from the area of optimal transport between probability measures. We demonstrate its effectiveness at aligning, combining and comparing time series living on incomparable spaces.

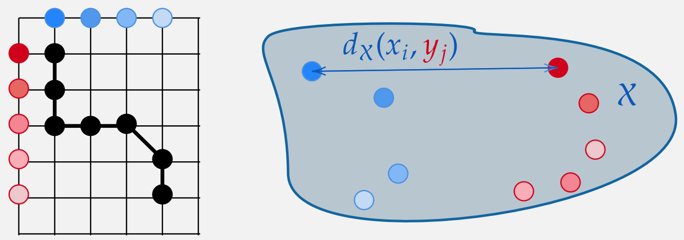
Dynamic Time Warping

Find alignment matrix minimizing total alignment cost

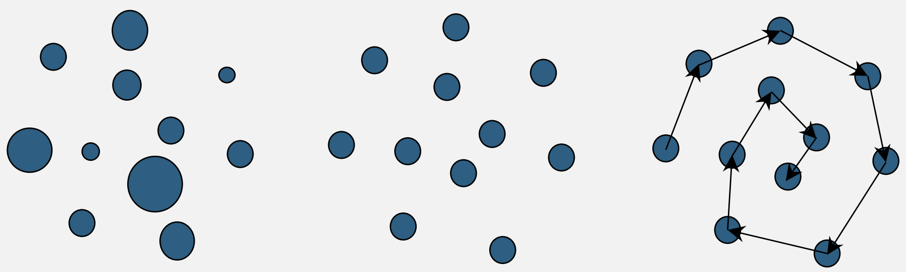
$$\mathrm{DTW}(\boldsymbol{x},\boldsymbol{y}) = \min_{\mathbf{A}\in\mathcal{A}(T_x,T_y)} \langle \mathbf{D}, \mathbf{A}\rangle_{\mathrm{F}} = \min_{\mathbf{A}\in\mathcal{A}(T_x,T_y)} \sum_{ij} d(x_i, y_j) A_{ij}$$

$\boldsymbol{x},\boldsymbol{y}$: time series of length T_x, T_y

$\mathbf{D}$: distance matrix $\quad$ $\mathbf{A}$: alignment matrix



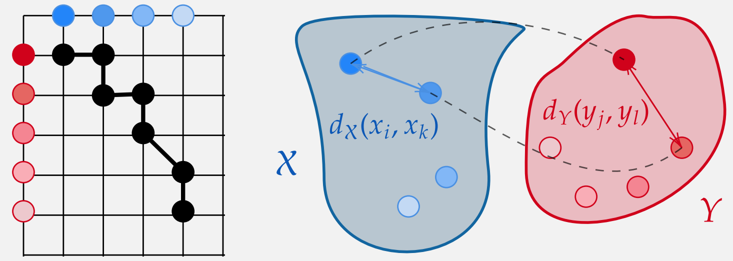
Connections with Optimal Transport



Time series ≈ ordered discrete measures

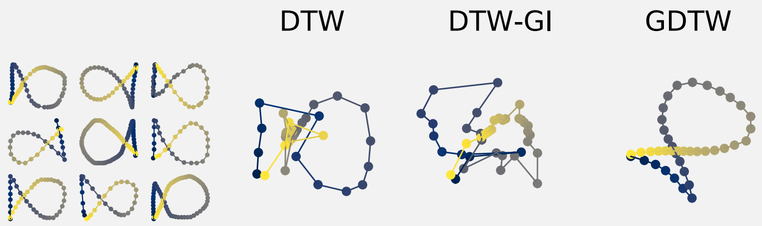
Gromov Dynamic Time Warping

$$\mathrm{GDTW}(\boldsymbol{x},\boldsymbol{y}) = \min_{\mathbf{A}\in\mathcal{A}(T_x,T_y)} \langle \mathbf{L}\otimes\mathbf{A}, \mathbf{A}\rangle_{\mathrm{F}}$$
$$= \min_{\mathbf{A}\in\mathcal{A}(T_x,T_y)} \sum_{ijkl} \mathcal{L}\Big[d_{\mathcal{X}}(x_i,x_k), d_{\mathcal{Y}}(y_j,y_l)\Big] A_{ij}A_{kl}$$



- ✓ Well-defined if x, y on different spaces
- ✓ Invariant under isometries by nature
- ✓ Estimated via Frank–Wolfe method

Application: Barycentric Averaging



Barycenter: any isometry class $\boldsymbol{x}$ satisfying

$$\mathbf{D}^* = \operatorname*{arg\,min}_{\mathbf{D}\in\mathbb{R}^{T\times T}} \sum_{j=1}^{J} \alpha_j \,\mathrm{GDTW}(\mathbf{D}, \mathbf{D}_{\boldsymbol{x}_j})$$

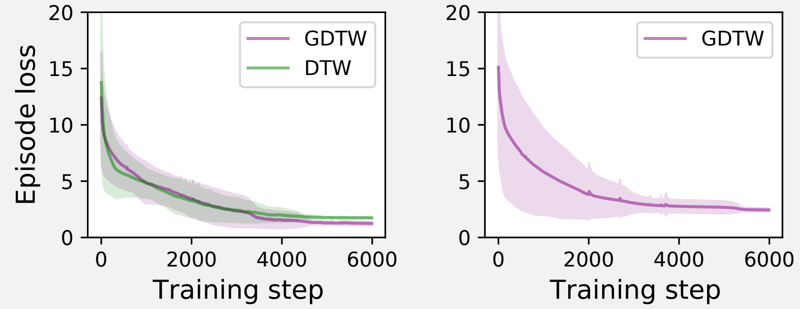
Application: Generative Modeling

Top: DTW samples $\quad$ Bottom: GDTW samples

$$\min_{\theta\in\Theta} \mathrm{W}_{\varepsilon}(\mu,\mu_\theta) := \min_{\pi\in\Pi(\mu,\mu_\theta)} \mathop{\mathbb{E}}_{(\boldsymbol{x},\boldsymbol{y})\sim\pi} \mathrm{GDTW}_\gamma(\boldsymbol{x},\boldsymbol{y}) - \varepsilon H(\pi),$$

$\boldsymbol{x}_j$	Time series data
μ	Sum of Dirac measures centered at data
μ_θ	Distribution of generative model
W_ε	Entropic Wasserstein with GDTW_γ ground cost

Application: Imitation Learning



Left: comparable spaces $\quad$ Right: incomparable spaces

$$\min_{\theta} \mathrm{GDTW}_\gamma(\boldsymbol{y}_{\mathrm{exp}}, \boldsymbol{x}_\theta).$$

$\boldsymbol{y}_{\mathrm{exp}}$	Expert state trajectory
$\boldsymbol{x}_\theta$	State trajectory under policy π_θ

References

[1] S. Cohen, G. Luise, A. Terenin, B. Amos, and M. P. Deisenroth. Aligning Time Series on Incomparable Spaces. AISTATS, 2021.
[2] H. Sakoe and S. Chiba. Dynamic Programming Algorithm Optimization for Spoken Word Recognition. ICASSP, 1978
[3] F. Zhou and F. Torre. Canonical Time Warping for Alignment of Human Behavior. NeurIPS, 2009
[4] T. Vayer, L. Chapel, N. Courty, R. Flamary, Y. Soullard, and R. Tavenard. Time Series Alignment with Global Invariances. arXiv, 2020.