Understanding the DWB codebase & Useage

Parameters specified in DataSetParameters.py
K, n ('window'), δ ('stride'), λ ('regObs'), η ('cyclicThresh'), s ('cluster_sig').

See comments in this file for more detail & references to appropriate equations in paper.

5 Model Estimation

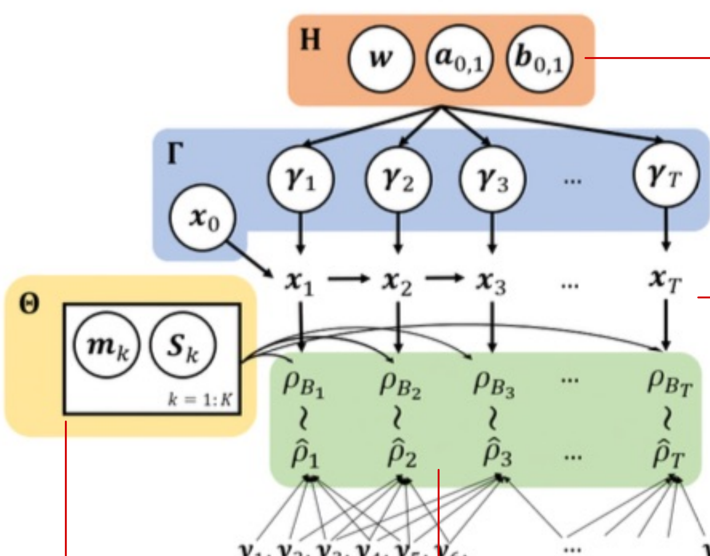
Algorithm 1: Dynamical Wasserstein Barycenter (DWB) Time-Series Estimation

Input:
$\boldsymbol{y}_\tau, \tau = 1 \dots \mathcal{T}$: Time series observations
K: Number of pure states

Hyperparameters:
n: Window size, δ: Window stride
λ: Weight on data-fit term
s: Variance on prior for $\boldsymbol{\Theta}$
learned in line 191 dwb.py (muP, covP) → (μ_0, σ_0): Mean, var. of $p(\boldsymbol{\Theta})$ reference dist.
η: Convergence threshold

Output:
$\boldsymbol{\Theta} = \left\{\{\boldsymbol{m}_k, \boldsymbol{S}_k\}_{k=1}^K\right\}$: Pure-state emission params
$\boldsymbol{\Gamma} = \left\{\boldsymbol{x}_0, \{\gamma_t\}_{t=1}^T\right\}$: Initial state and innovations
$\boldsymbol{X} = \{\boldsymbol{x}_t\}_{t=1}^T$: Wasserstein barycentric state vector (Computed from $\boldsymbol{\Gamma}$ via (7))
$\boldsymbol{H} = \{\boldsymbol{w}, \boldsymbol{a}_1, \boldsymbol{b}_1\}$: Beta mixture parameters for transition dynamics

for $t = 1, \dots, T$ where $T = \lfloor \frac{(\mathcal{T}-(2n+1))}{\delta} \rfloor + 1$ **do**
  $\boldsymbol{m}_t = \frac{1}{(2n+1)} \sum_{i=1}^{2n+1} \boldsymbol{y}_{\delta(t-1)+i}$; `// Preprocessing of windowed`
  $\boldsymbol{S}_t = \frac{1}{2n} \sum_{i=1}^{2n+1} (\boldsymbol{y}_{\delta(t-1)+i} - \boldsymbol{m}_t)(\boldsymbol{y}_{\delta(t-1)+i} - \boldsymbol{m}_t)^T$; `// empirical distributions`
  $\hat{\rho}_t = \mathcal{N}(\boldsymbol{m}_t, \boldsymbol{S}_t)$
end
$c^{(0)} = F\left(\boldsymbol{\Gamma}^{(0)}, \boldsymbol{\Theta}^{(0)}, \boldsymbol{H}^{(0)}, \{\hat{\rho}_t\}_{t=1}^T\right)$; `// Cost function F defined in (9)`
do
  $\boldsymbol{\Gamma}^{(n+1)}, \boldsymbol{H}^{(n+1)} = \operatorname{argmin}_{\boldsymbol{\Gamma}, \boldsymbol{H}} F\left(\boldsymbol{\Gamma}^{(n)}, \boldsymbol{\Theta}^{(n)}, \boldsymbol{H}^{(n)}, \{\hat{\rho}_t\}_{t=1}^T\right)$; `// Adam`
  $\boldsymbol{\Theta}^{(n+1)} = \operatorname{argmin}_{\boldsymbol{\Theta}} F\left(\boldsymbol{\Gamma}^{(n+1)}, \boldsymbol{\Theta}^{(n)}, \boldsymbol{H}^{(n+1)}, \{\hat{\rho}_t\}_{t=1}^T\right)$; `// Riemannian line search`
  $c^{(n+1)} = F\left(\boldsymbol{\Gamma}^{(n+1)}, \boldsymbol{\Theta}^{(n+1)}, \boldsymbol{H}^{(n+1)}, \{\hat{\rho}_t\}_{t=1}^T\right)$
while $(c^{(n)} - c^{(n+1)}) > \eta$;



Sec 4.1 in paper — (learned) hyperparameters for the innov. vectors (Beta dist)

Sec 4.2 in paper — Innovation vectors which determines the state vectors

sec 3 in paper — State vector indexed by sliding window

eqns (6) & (9) in paper + sec 4.2 for prior on Θ — Parameters (learned) for pure state gaussians

eqn (3) for weighted Barycentre eqn & sec 3 for empirical estimation — Model-predicted Wasserstein Barycentres of pure states according to state vector vs. empirically estimated Gaussians