"Dimensionality Reduction for Wasserstein Barycenter"

Zachary Izzo, Sandeep Silwal and Samson Zhou, October 2021

Gonzague de Carpentier
Professor: Gabriel Peyré

January 2023

Introduction

Discrete Wasserstein barycenter problem:

- Given:
  - Discrete measures: $\mu^{(i)} = \sum_{j=1}^{T} a_j^{(i)} \delta_{x_j^{(i)}}$ where $x_j^{(i)} \in \mathbb{R}^d$
- Compute:

$$\nu^* \in \arg\min_{\nu \in \mathcal{P}} \sum_{i=1}^{k} W_p(\mu^{(i)}, \nu)^p$$

where $\mathcal{P} = \left\{ \sum_{j=1}^{n} b_j \delta_{y_j}, b_j \geqslant 0, y_j \in \mathbb{R}^d \right\}$

Issue: computation of discrete WB is NP-hard (runtime exponential in dimension) [2]

Proposed approach in [3] : dimensionality reduction

Characterization of the Wasserstein Barycenter

Let $w^{*(i)}$ be the optimal coupling between $\mu^{(i)}$ and ν^*. Then,

$$b^*_{j'} = \sum_j w^{*(i)}_{j,j'} \tag{1}$$

$$y^*_{j'} = \arg\min_y \sum_{i,j} w^{*(i)}_{j,j'} ||x^{(i)}_j - y_{j'}||^p \tag{2}$$

From now on $p = 2$.

Johnson-Lindenstrauss projections

Let $X \subset \mathbb{R}^d$ be a set of n points. Let $A \in \mathbb{R}^{m \times d}$ be a random matrix whose entries are sampled independently from $\mathcal{N}(0, 1/m)$. Let $\varepsilon \in (0, 1/2)$. Then,

$$\mathbb{P}(\forall u, v \in X, (1-\varepsilon)||u-v||^2 \leqslant ||Au - Av||^2 \leqslant (1+\varepsilon)||u-v||^2) \\ \geqslant 1 - 2n^2 e^{-\varepsilon^2 m/8}$$

All distances are preserved up to a factor $1 \pm \varepsilon$ with high probability ($\geqslant 95\%$) if $m \geqslant \frac{16}{\varepsilon^2}(2 + \log(n))$. [4]

Proposed method

The authors propose following dimensionality reduction method:

1. Project measures $\mu^{(i)}$ into $\mathbb{R}^m$ using the projection A: $\tilde{\mu}^{(i)} = \sum_{j=1}^{T} a_j^{(i)} \delta_{Ax_j^{(i)}}$
2. Compute the Wasserstein barycenter $\tilde{\nu}^*$ of the $(\tilde{\mu}^{(i)})$ and associated weights $\tilde{b}^*$, support points $\tilde{y}^*$ and optimal couplings $\tilde{w}^*$
3. Compute an approximation of the Wasserstein barycenter in $\mathbb{R}^d$ by computing b^* and y^* from $w^* := \tilde{w}^*$

Numerics

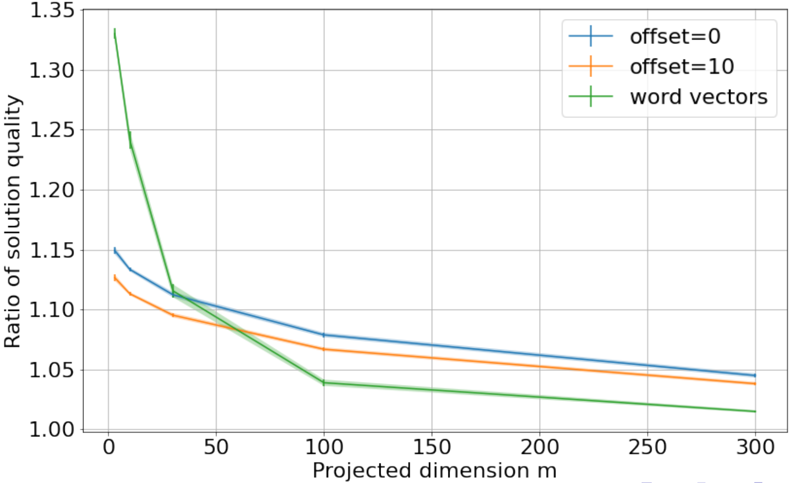
Experiments with $k = 2$ measures $\rightarrow$ WB = interpolation of the optimal transportation plans computed with **ot.emd** from the Python Optimal Transport library.

3 types of pairs of measures $(\mu^{(1)}, \mu^{(2)})$:

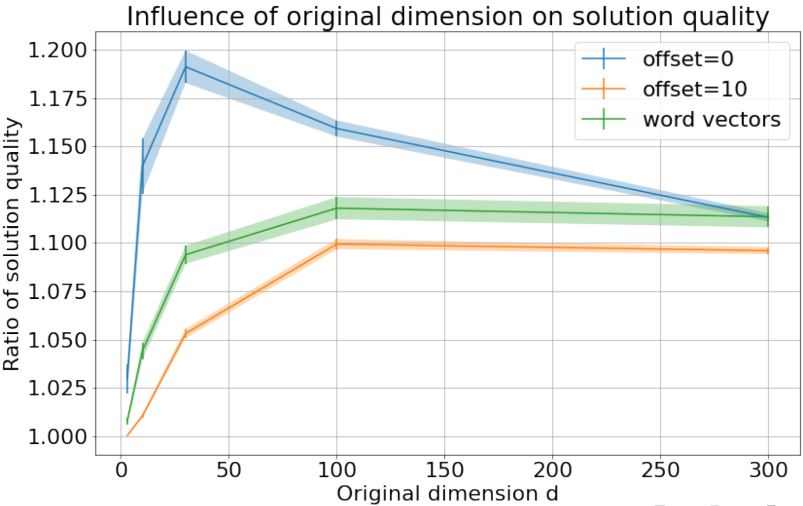
- Uniform weights $a_j^{(i)} = 1/T$ and $x_j^{(i)}$ sampled randomly and independently from $\mathcal{N}(0, I)$ in $\mathbb{R}^{300}$
- Same as before but with offset of 10 on the first dimension for $\mu^{(2)}$
- Aligned word vectors for French and English in dimension 300 downloaded from fastText [1].

$$\text{Quality ratio} = \frac{\text{cost of the WB computed with DR}}{\text{cost of the true WB}}$$

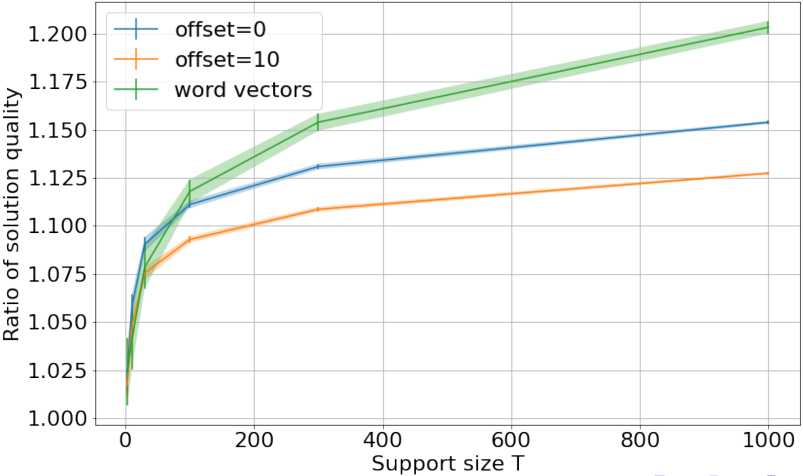
Influence of the projected dimension m on the solution quality ($d = 300,\ T = n = 100$)



Influence of the original dimension d on the solution quality ($m = 30,\ T = n = 100$)



Influence of the support size T on the solution quality ($d = 300$, $m = 30$)

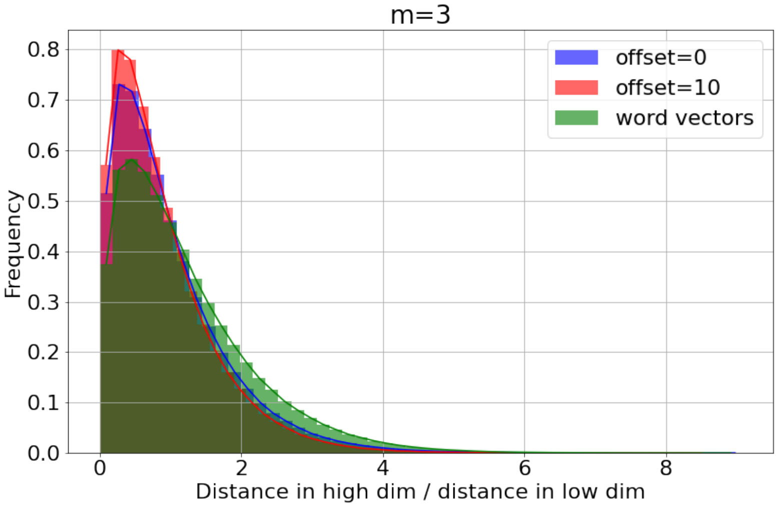


Distance matrix distortion

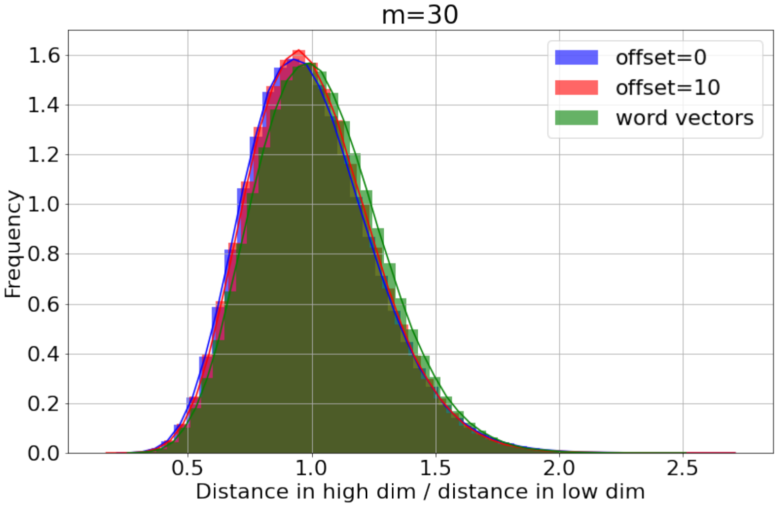
For two measures $(\mu^{(1)}, \mu^{(2)})$:

- $D_{j,j'} := ||x_j^{(1)} - x_{j'}^{(2)}||^2$
- $\tilde{D}_{j,j'} := ||Ax_j^{(1)} - Ax_{j'}^{(2)}||^2$
- Plot histogram of values $\tilde{D}_{j,j'}/D_{j,j'}$

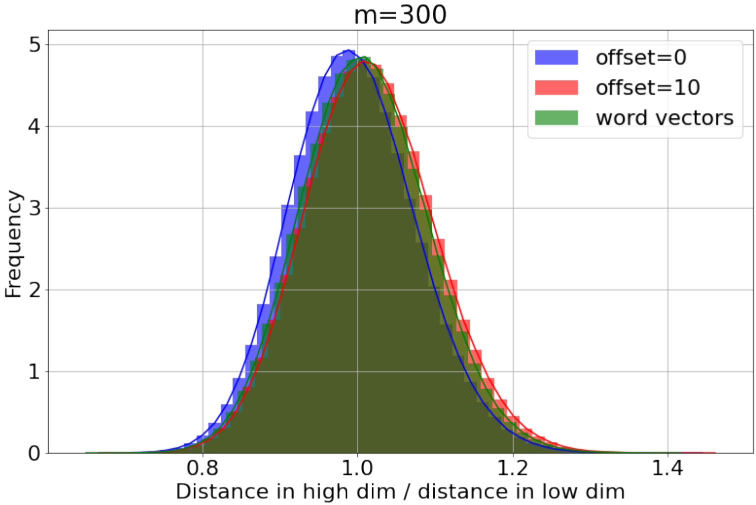
Distance matrix distortion ($d = 300,\ T = n = 1000$)



Distance matrix distortion ($d = 300,\ T = n = 1000$)



Distance matrix distortion ($d = 300,\ T = n = 1000$)



Conclusion

- Powerful method: get rid of the exponential dependence in the dimension.
- However: projected dimension proportional to $1/\varepsilon^2$.

References

[1] *Aligned word vectors · fastText*. URL: https://fasttext.cc/index.html (visited on 01/04/2023).

[2] Jason M. Altschuler and Enric Boix-Adsera. *Wasserstein barycenters can be computed in polynomial time in fixed dimension*. Dec. 9, 2020. arXiv: 2006.08012[cs,math]. URL: http://arxiv.org/abs/2006.08012 (visited on 12/26/2022).

[3] Zachary Izzo, Sandeep Silwal, and Samson Zhou. *Dimensionality Reduction for Wasserstein Barycenter*. Oct. 18, 2021. arXiv: 2110.08991[cs,math]. URL: http://arxiv.org/abs/2110.08991 (visited on 12/26/2022).

[4] Sham Kakade and Greg Shakhnarovich. "Random Projections". In: (2009).