MIIS 2022
International Workshop On Mathematical Issues In Information Sciences
数据科学国际研讨会
17th-18th, Dec. 2022

# Metric Nearness Made Practical

Wenye Li[1,2], Fangchen Yu[1], Zichen Ma[1]
[1] The Chinese University of Hong Kong, Shenzhen
[2] Shenzhen Research Institute of Big Data

## Introduction

Consider the Metric Nearness Problem [1]:

$$\min_{X \in \mathbb{R}^{n \times n}} \|X - D^o\|_F^2$$

$$\text{subject to} \begin{cases} x_{ii} = 0 \\ x_{ij} = x_{ji} \geq 0 \quad , \quad \forall 1 \leq i,j,k \leq n \\ \boldsymbol{x_{ij} \leq x_{ik} + x_{kj}} \end{cases}$$

- Metric nearness model seeks a valid metric $X$ that is nearest to the observed non-metric $D^o$.
- In practice, existing approaches still face non-trivial challenges from a large number of $\boldsymbol{O(n^3)}$ **constraints**.

## Proposed Method

We designed a two-stage approach to solve it.

### Stage-I. Embedding Calibration

The approach first shrinks the scope of distance metrics to *isometrically embeddable matrices*.

- Schoenberg's result on isometrical embedding provides a sufficient and necessary condition [2].

**Theorem 1.** *$D = \{d_{ij}\} \in \mathbb{R}^{n \times n}$ is an embeddable matrix iff the matrix $E = \exp(-\gamma D)$ is PSD for $\gamma > 0$.*

- We seek an embeddable matrix by solving.

$$E^* = \min_{E \in \mathbb{R}^{n \times n}} \|E - E^o\|_F^2$$

$$\text{subject to} \begin{cases} e_{ii} = 1, 0 \leq e_{ij} = e_{ji} \leq 1, \forall 1 \leq i,j \leq n \\ E \succcurlyeq 0 \text{ (PSD)} \end{cases}$$

Dykstra's projection algorithm [3] is conducted on two closed convex sets, i.e., $\mathcal{S}$ and $\mathcal{T}$, defined by

$\mathcal{S} = \{E \in \mathbb{R}^{n \times n} | E \succcurlyeq 0\}$,

$\mathcal{T} = \{E \in \mathbb{R}^{n \times n} | e_{ii} = 1, 0 \leq e_{ij} = e_{ji} \leq 1, \forall i,j\}$.

where running for a few iterations is efficient to obtain a good estimate $X^0 = -\log(E^*)/\gamma$.

### Stage-II. HLWB Projection

Then the approach starts with the initial solution in Stage-I and refines it iteratively to the optimum.

- The region defined by the triangle inequalities is the intersection of all $C_{ijk}$'s, denoted by

$$C_{ijk} = \{X \in \mathbb{R}^{n \times n} | x_{ij} \leq x_{ik} + x_{kj}\}.$$

- We use a HLWB projection [4] to sequentially project a given point $D^o$ onto the mulitple closed convex sets.

**Theorem 2.** *Let $\mathcal{C}_1, \cdots, \mathcal{C}_m$ be a family of closed convex subsets such that $\mathcal{M}_n = \cap_{i=1}^m \mathcal{C}_i \neq \emptyset$. Set*

$$\begin{cases} Y^{t+1} = \frac{1}{t+2} D^o + \frac{t+1}{t+2} X^t \\ X^{t+1} = \mathrm{P}_{\mathcal{C}_1} \cdots \mathrm{P}_{\mathcal{C}_m}(Y^t) \end{cases}, \text{for } t = 0, 1, \cdots.$$

*Then $X^t \to \mathrm{P}_{\mathcal{M}_n}(D^o), Y^t \to \mathrm{P}_{\mathcal{M}_n}(D^o)$ as $t \to \infty$.*

## Algorithm

The algorithm has $O(n^3)$ time & $O(n^2)$ space complexity.

**Algorithm 1:** The Proposed HLWB Algorithm

```
X^0 ← Calibrate(D^o)                        ▷ Step 1: Embedding Calibration
for t ← 1, ···, maxiter do                   ▷ Step 2: HLWB Projection
    X^t ← 1/(t+2) × D^o + (t+1)/(t+2) × X^{t-1};
    for each (i, j, k) do
        δ ← (X^t_ij − X^t_ik − X^t_kj)/3;
        if δ > 0 then                        ▷ violation detected
            X^t_ij ← X^t_ij − δ;
            X^t_ik ← X^t_ik + δ;
            X^t_kj ← X^t_kj + δ;
```

## Evaluation

### Result-I. Problem Size

**Table 1.** The largest problem size $n$ solved within 12 hours.

| CPLEX | MOSEK | TRF [1] | PAF [5] | **HLWB** |
|---|---|---|---|---|
| $< 300$ | $< 300$ | $< 2,000$ | $\sim 3,000$ | $> 10,000$ |

### Result-II. Nearness and Optimality

Nearness is measured by $NMSE = \|X^* - D^o\|_F^2 / \|D^o\|_F^2$.

Constraint Satisfaction Ratio $CSR = \#satisfied/\#total$.

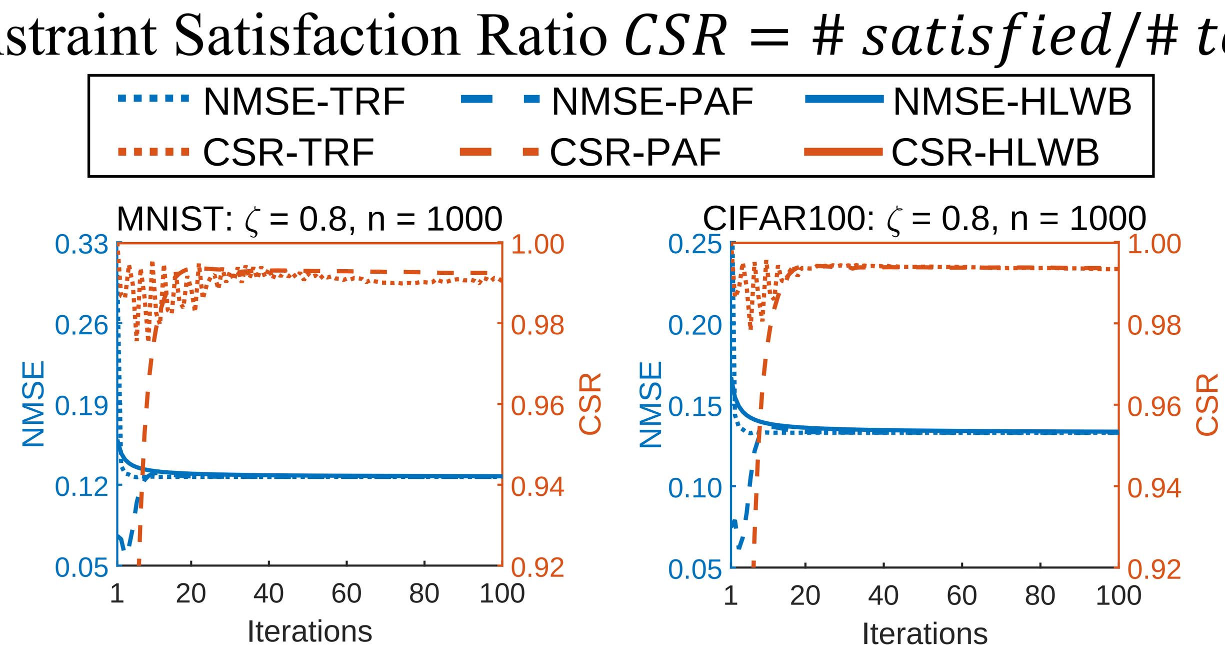


**Figure 1.** NMSE/CSR vs Iterations on Noisy Distance.

$D^o = \{d_{ij}^o\} = \{\max\{0, d_{ij}^* + \zeta \cdot \mathrm{mean}(D^*) \cdot \mathrm{N}(0,1)\}\}$.

### Result-III. Updates and Running Time

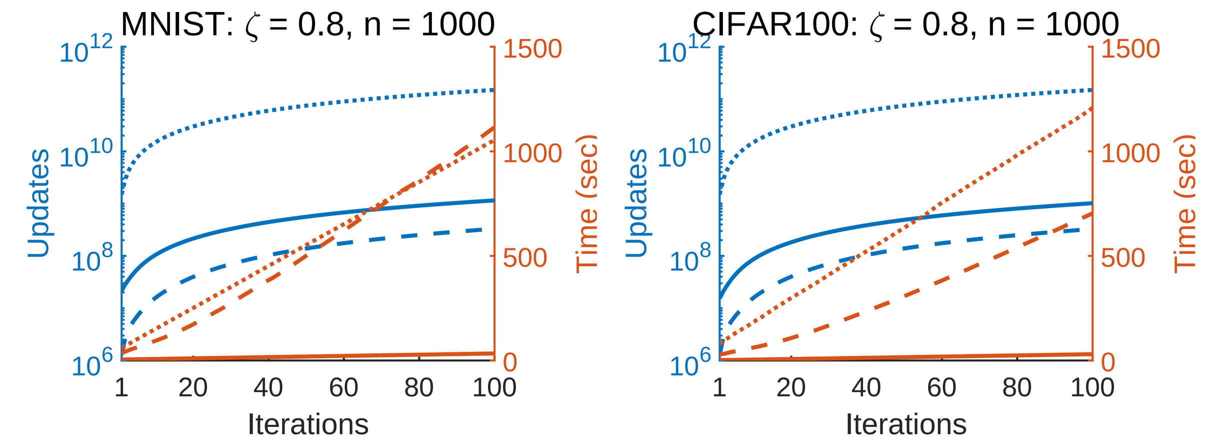


**Figure 2.** Updates/Time vs Iterations on Noisy Distance.

[1] Brickell, J., etc. 2008. The metric nearness problem. *SIAM Journal on Matrix Analysis and Applications*, 30(1): 375–396.
[2] Schoenberg, I. J. 1938. Metric spaces and positive definite functions. *Transactions of the American Mathematical Society*, 44(3): 522–536.
[3] Boyle, J. P., etc. 1986. A method for finding projections onto the intersection of convex sets in Hilbert spaces. *In Advances in Order Restricted Statistical Inference*, 28–47. Springer.
[4] Halpern, B. 1967. Fixed points of nonexpanding maps. *Bulletin of the American Mathematical Society*, 73(6): 957–961.
[5] Sonthalia, R., etc. 2020. Project and forget: Solving large-scale metric constrained problems. *arXiv preprint arXiv:2005.03853*.