Multi-Criteria Dimensionality Reduction with Applications to Fairness

UTHAIPON (TAO) TANTIPONGPIPAT

(GRADUATING) PHD STUDENT, INDUSTRY JOB MARKET

GEORGIA INSTITUTE OF TECHNOLOGY

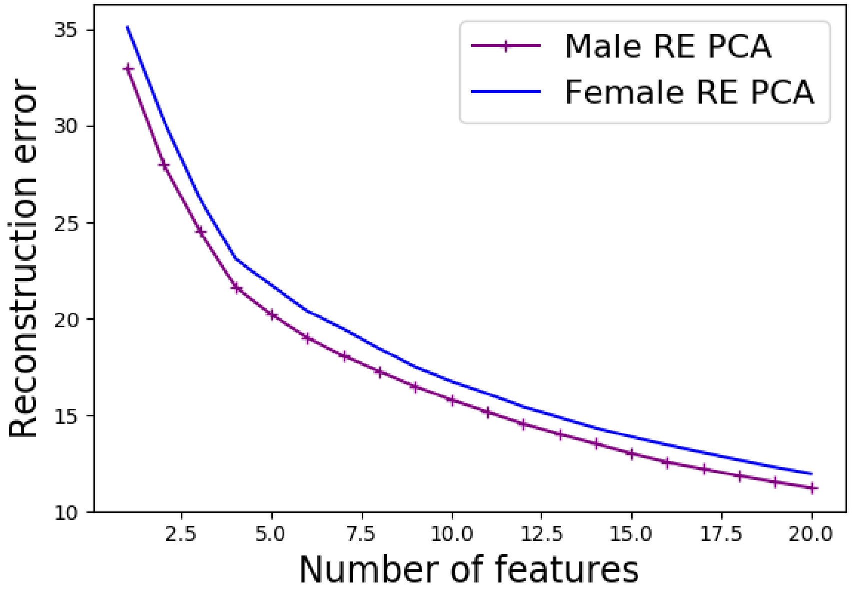
JOINT WORK WITH

JAMIE MORGERNSTERN, SAMIRA SAMADI, MOHIT SINGH, AND SANTOSH VEMPALA

PCA can be unfair!

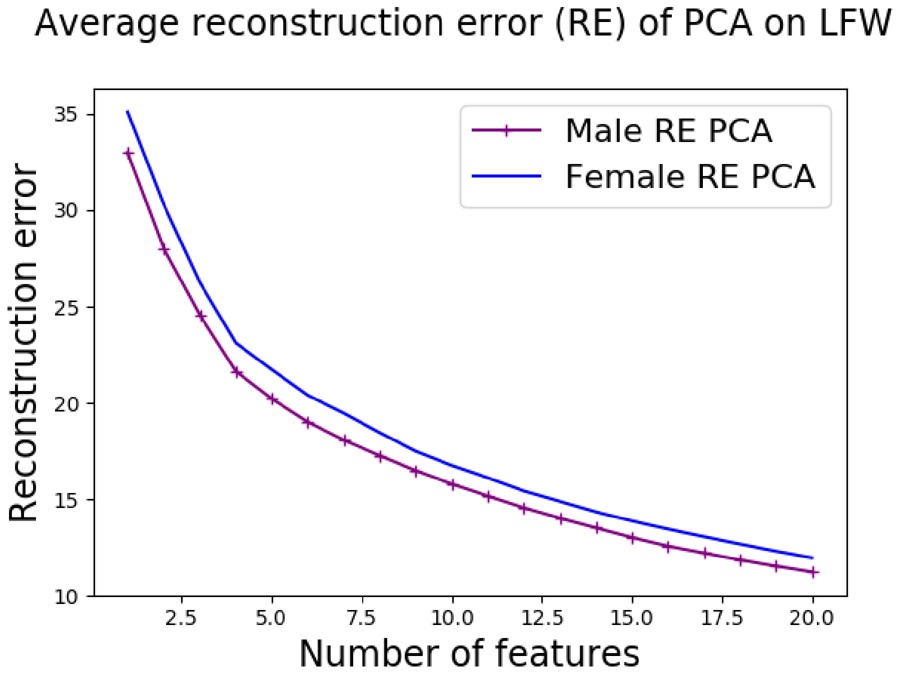
Standard PCA on face data LFW of male and female

Average reconstruction error (RE) of PCA on LFW

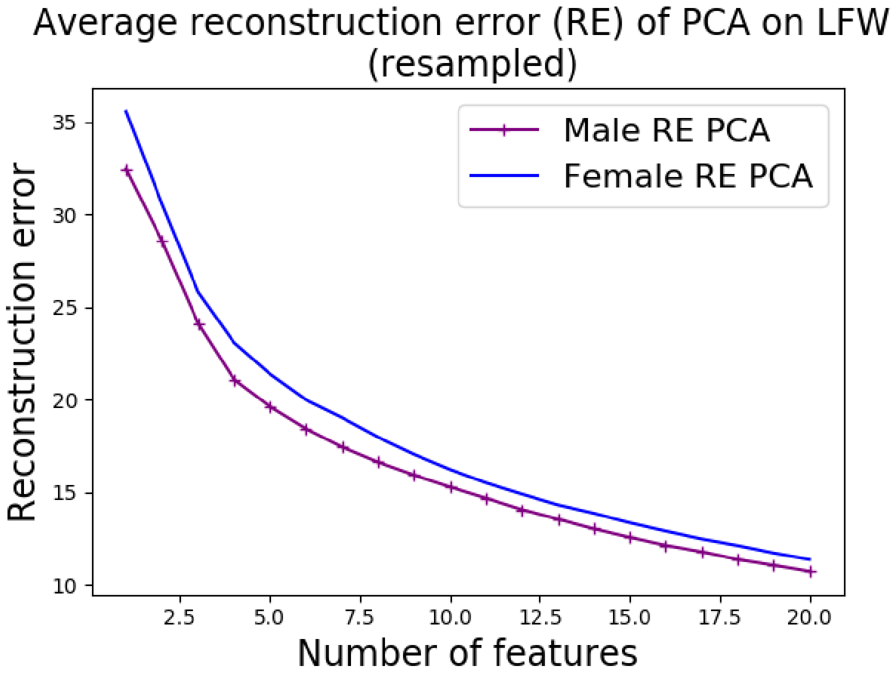


PCA can be unfair!

Standard PCA on face data LFW of male and female



Equalizing male and female weight before PCA



Contribution 1: Problem Formulation

Multi-criteria dimensionality reduction (MCDR):

$$\max_{\text{projection } P} \; g\big(f_1(P), f_2(P), \ldots, f_k(P)\big)$$

Utility criterion f_i's and social welfare g

Contribution 1: Problem Formulation

Multi-criteria dimensionality reduction (MCDR):

$$\max_{\text{projection } P} \; g\big(f_1(P), f_2(P), \ldots, f_k(P)\big)$$

Utility criterion f_i's and social welfare g

- *Mar-Loss*: $\min_P \max_{i\in\{1,\ldots,k\}} \left(\max_Q \|A_i Q\|_F^2 - \|A_i P\|_F^2 \right)$
- *NSW*: $\max_P \quad \prod_{i=1}^{k} \|A_i P\|_F^2$

Contribution 2: Algorithms and Guarantees

On linear f_i in PP^{T} and concave g:

- Polynomial-time algorithm for MCDR with optimal utility and small rank violation $s = \sqrt{2k + 1/4} - 3/2$
- Approximation ratio $1 - s/d$ on utility when no rank violation

Contribution 2: Algorithms and Guarantees

On linear f_i in PP^{T} and concave g:

- Polynomial-time algorithm for MCDR with optimal utility and small rank violation $s = \sqrt{2k + 1/4} - 3/2$
- Approximation ratio $1 - s/d$ on utility when no rank violation
- Semi-definite Program (SDP) → Multiplicative Weight (MW) method
  - scalable up to ≈ 1000 dimensions

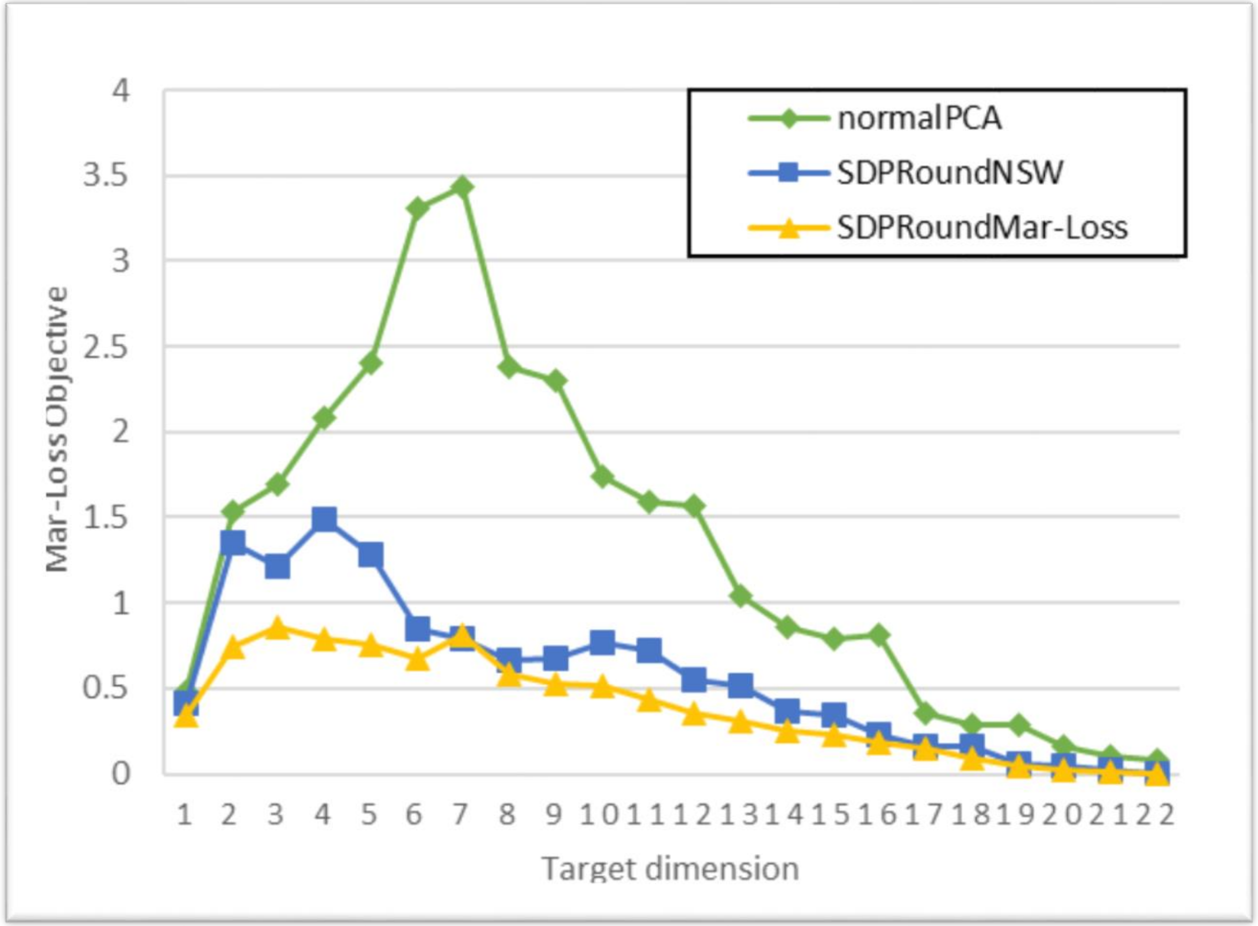
Contribution 2: Algorithms and Guarantees

- *Mar-Loss*:

$$\min_{P} \max_{i\in\{1,\ldots,k\}} \left(\max_{Q} \|A_i Q\|_F^2 - \|A_i P\|_F^2 \right)$$

- *NSW*:

$$\max_{P} \quad \prod_{i=1}^{k} \|A_i P\|_F^2$$



Contribution 3: Optimization Theory

- Every extreme point of the semi-definite program relaxation of MCDR has low rank
  - Generalize work on low-rank property in semi-definite program by Barvinok'95, Pataki'98
- Optimization result + ML application

Contribution 4: Complexity of MCDR

- NP-hard for general k
  - Reduction to MAX-CUT
- Polynomial-time for fixed k
  - Algorithmic theory of quadratic maps.

More details

- Poster: Thursday Dec 12th at 10:45 AM -- 12:45 PM, East Exhibition Hall B + C #80
- Happy to chat!

- Code: github.com/uthaipon/multi-criteria-dimensionality-reduction
- Web: sites.google.com/site/ssamadi/fair-pca-homepage

(searchable links at NeurIPS website or my website Uthaipon Tantipongpipat)