CS 754 : Project Report

Estimation of the sample covariance matrix from compressive measurements

Original paper by Farhad Pourkamali-Anaraki (2016)

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\Sigma = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)^t$$

Team Members :
Abhijat Bharadwaj (210020002)
Animesh Kumar (21D070012)

Overview

Objective:

- Given: A compressive measurement of a large data matrix is available
- To-Do: Estimate the sample covariance matrix
- To-Do: Analyze and improve accuracy of estimation

Need for Estimation:

- Importance of sample covariance matrix (PCA, KL-Transform)
- PCA needs eigen-decomposition of covariance matrix
- Complete data matrix often unavailable

Overview

Estimator 1: *The Biased Estimator*

$$\widehat{\mathbf{C}}_n := \frac{1}{(m^2+m)\mu_2^2} \cdot \frac{1}{n} \sum_{i=1}^{n} \mathbf{R}_i \mathbf{R}_i^T \mathbf{x}_i \mathbf{x}_i^T \mathbf{R}_i \mathbf{R}_i^T$$

Estimator 2: *The (New and Improved) Unbiased Estimator*

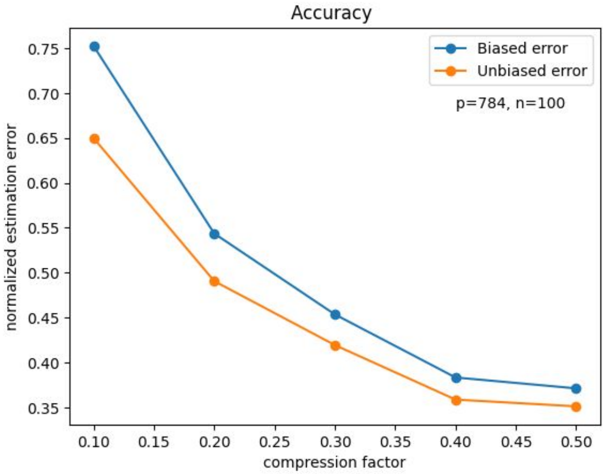
$$\widehat{\mathbf{\Sigma}}_n := \widehat{\mathbf{C}}_n - \alpha_1 \operatorname{diag}\left(\widehat{\mathbf{C}}_n\right) - \alpha_2 \operatorname{tr}\left(\widehat{\mathbf{C}}_n\right) \mathbf{I}_{p \times p}$$

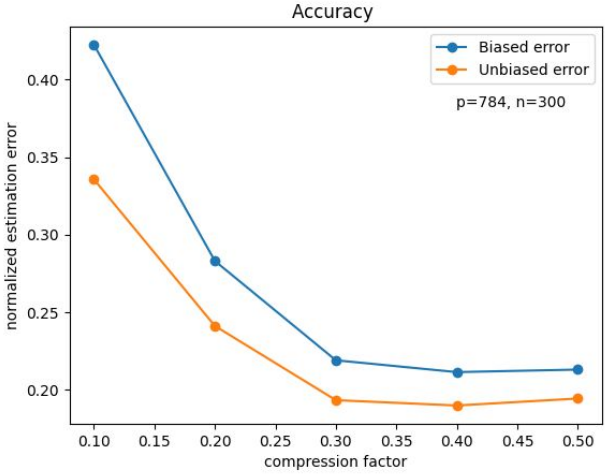
$$\alpha_1 := \frac{\frac{\kappa}{m+1}}{1 + \frac{\kappa}{m+1}}$$

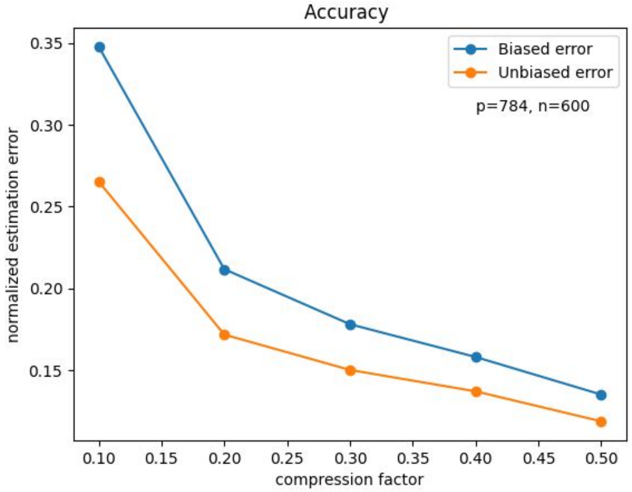
$$\alpha_2 := \frac{1}{(1 + \frac{\kappa}{m+1})(m+1+\kappa+p)}.$$

Results

MNIST Dataset :

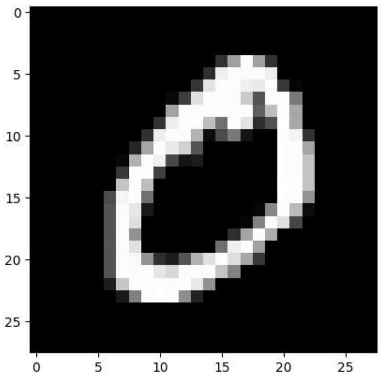


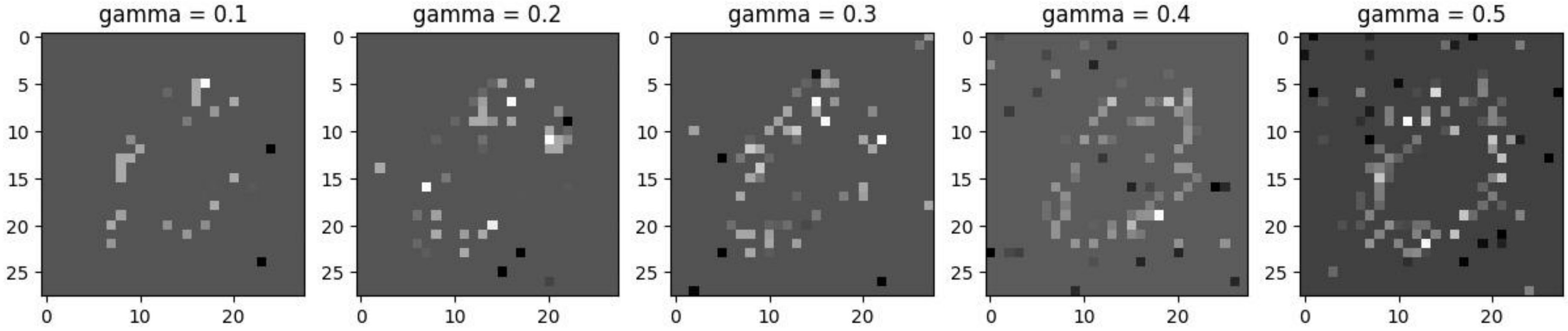




Results

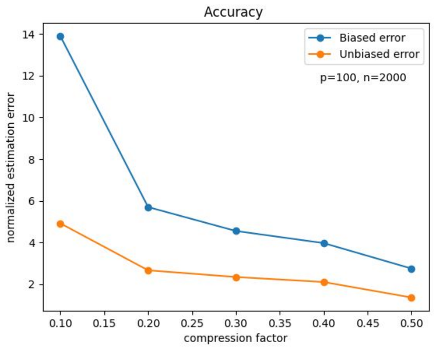
MNIST Dataset :

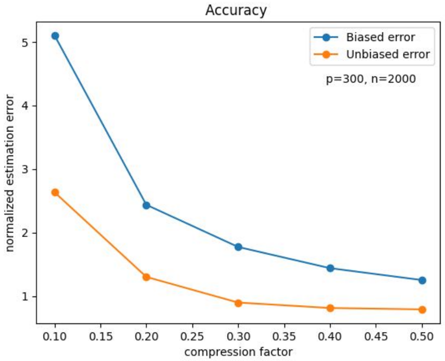


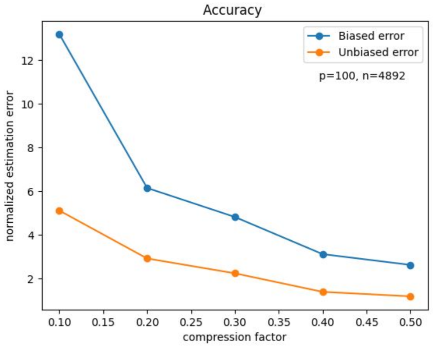


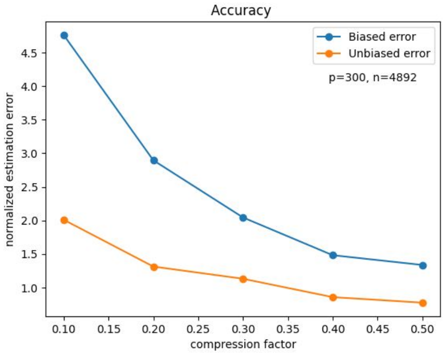
Results

Gen4 Dataset :



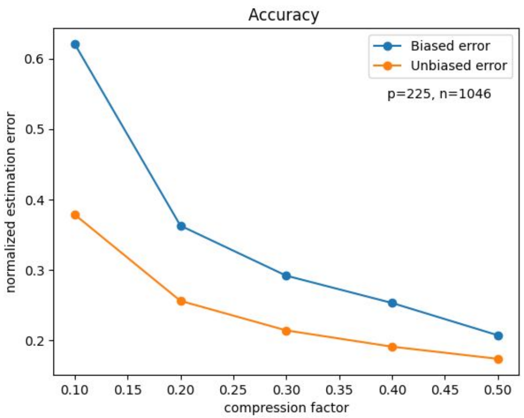


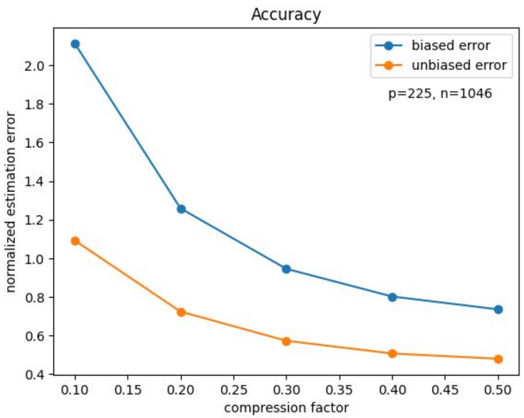


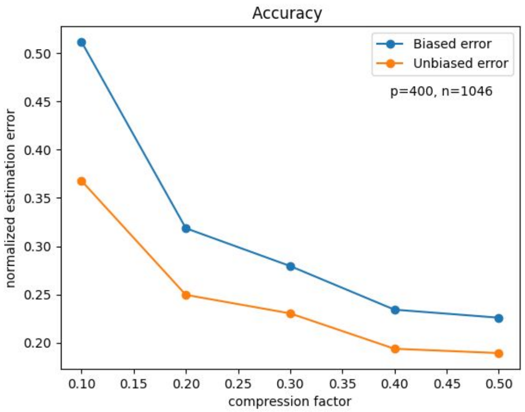


Results

Traffic Dataset :

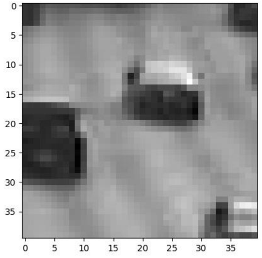


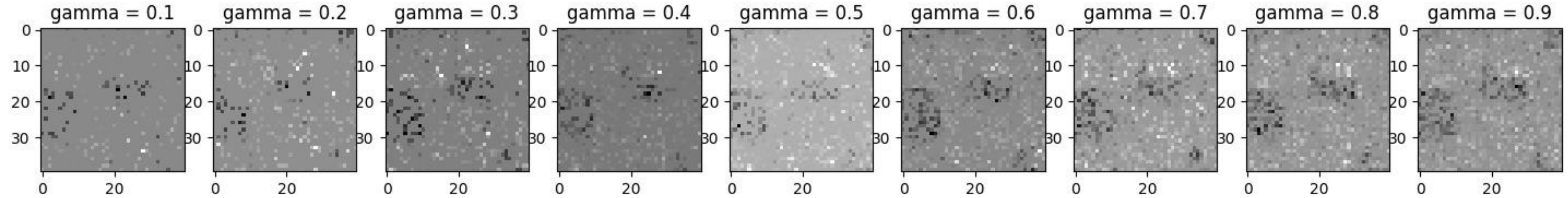




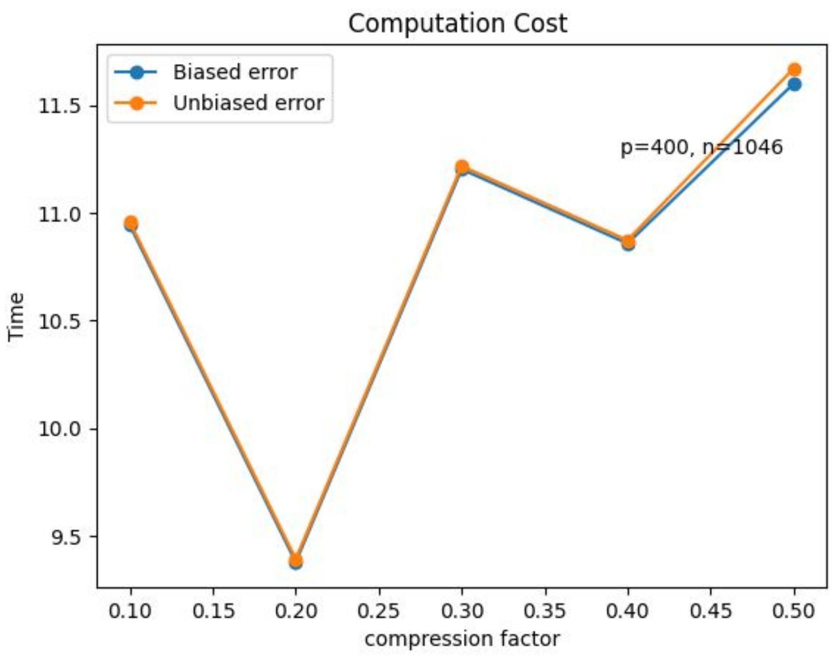
Results

Traffic Dataset :

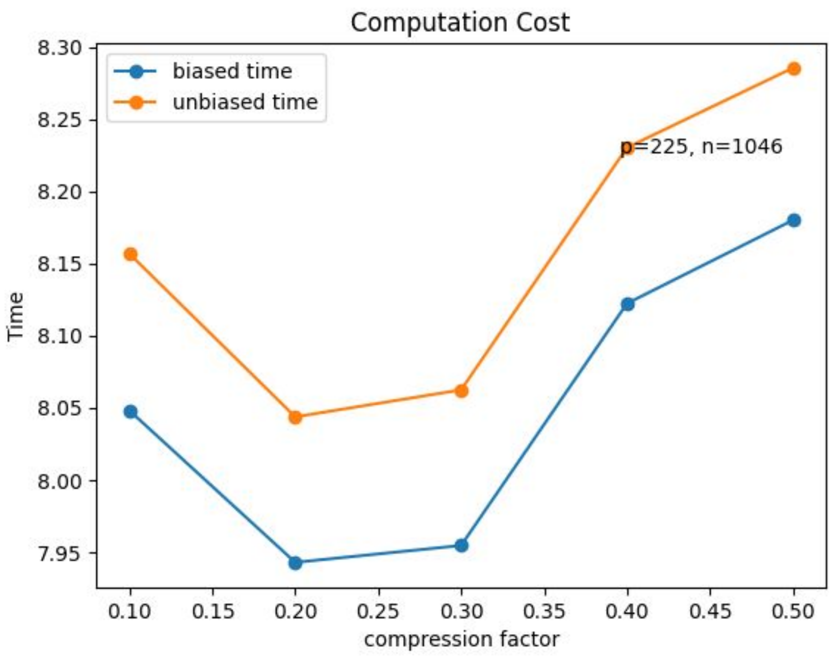




Results

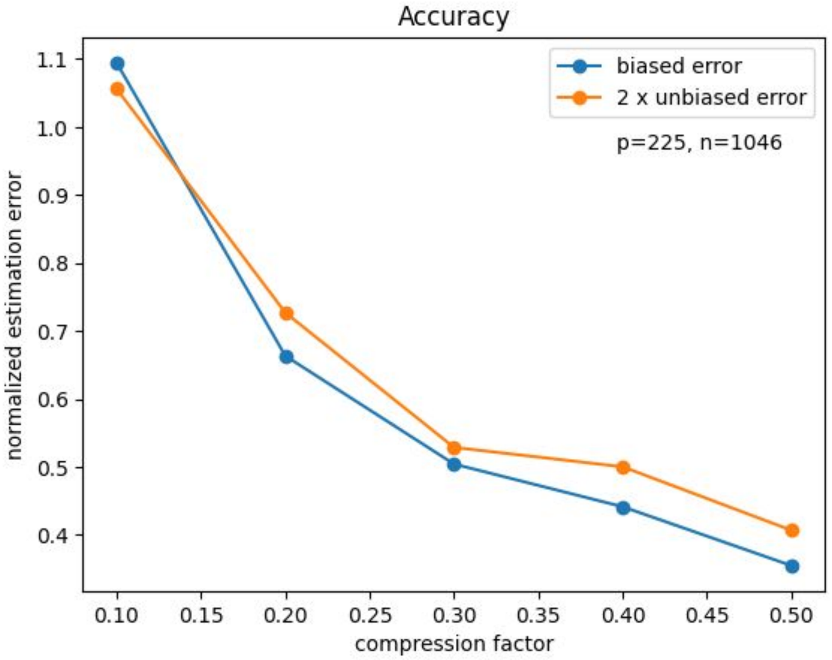


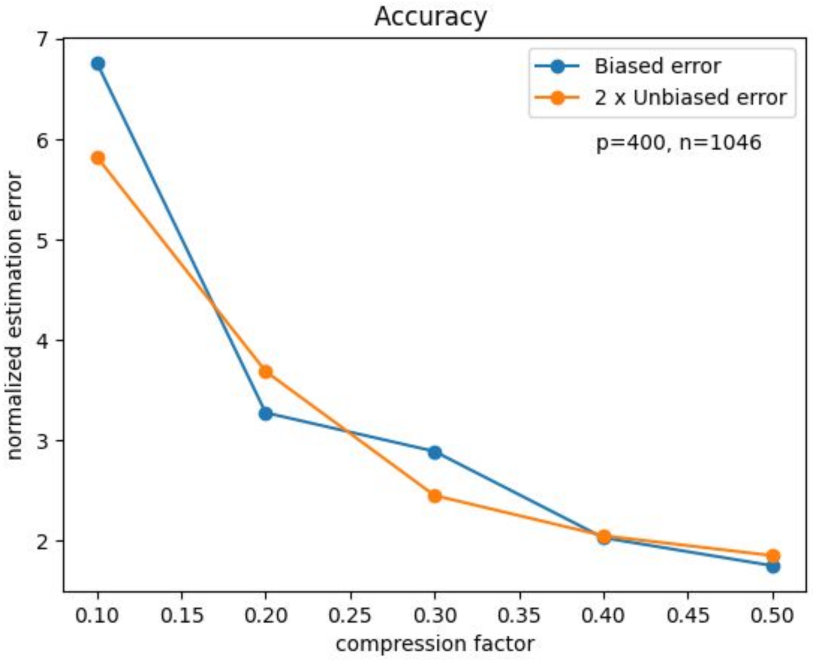
Computation Cost
Biased error
Unbiased error
p=400, n=1046
11.5
11.0
10.5
10.0
9.5
Time
0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45 0.50
compression factor



Computation Cost
biased time
unbiased time
p=225, n=1046
8.30
8.25
8.20
8.15
8.10
8.05
8.00
7.95
Time
0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45 0.50
compression factor

Results


Accuracy
biased error
2 x unbiased error
p=225, n=1046
normalized estimation error
compression factor
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50


Accuracy
Biased error
2 x Unbiased error
p=400, n=1046
normalized estimation error
compression factor
7
6
5
4
3
2
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50

Inferences

1. Estimation error decreases with increase in compression factor(γ)

2. Increasing the number of samples (n) decreases the error for a constant feature size (p) ***(see results from MNIST and Gen4)***

3. Feature size (p) affects the error more than the number of samples (n) ***(see results from Gen4)***

4. Unbiased Error ~ ½ * Biased Error