Outline

Evaluation Criteria:

1. Proportion of Variance Explained by the first k factors/principal components.

$$\mathrm{PVE} = \frac{\mathrm{trace}(X_k^T X_k)}{\mathrm{trace}(X^T X)}, \text{ where } X_k = X V_k (X_k^T V_k)^{-1} V_k^T$$

2. Percentage of Error Unexplained:

$$\mathrm{PEU} = \frac{\|\Sigma - \widehat{\Sigma}\|_F}{\|\Sigma\|_F}, \text{ where } \widehat{\Sigma} = \frac{\widehat{V}\widehat{D}\widehat{U^T}\widehat{U}\widehat{D}\widehat{V^T}}{n}$$

Case 1: Dense Eigenvector

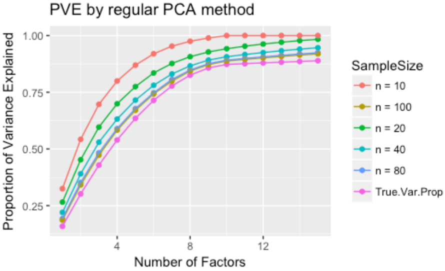
Simulation Settings:

- Number of variables, $p = 50$
- Eigenvalues:

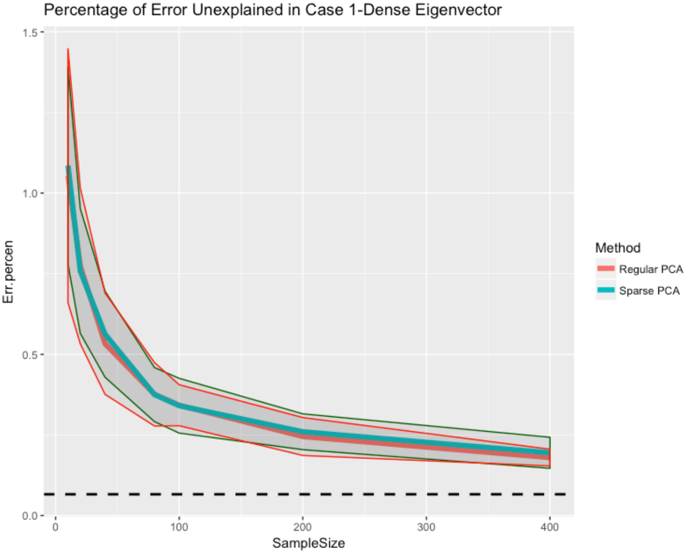
$$(\lambda_1, \lambda_2, \ldots, \lambda_{10}, \lambda_{11}, \ldots, \lambda_{50}) = (\underbrace{5, 4.5, \ldots, 0.5}_{10}, \underbrace{0.1, 0.1, \ldots, 0.1}_{40})$$

- Eigenvectors: 50 orthonormal vectors, elements in eigenvectors are non-zero.
- Sample Size, $n = 10, 20, 40, 80, 100$
- Number of repetitions, $np = 20$

Case 2: Sparse Eigenvector, PVE



Case 1: Dense Eigenvector, PEU



Case 2: Sparse Eigenvector

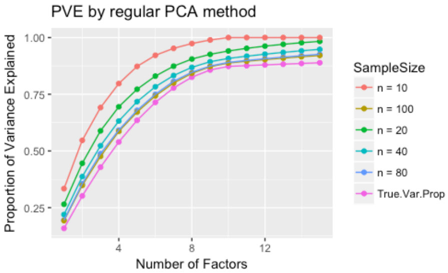
Simulation Settings:

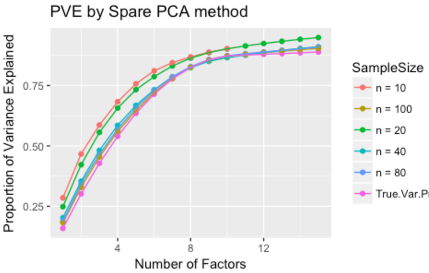
- Number of variables, $p = 50$
- Eigenvalues:

$$(\lambda_1, \lambda_2, \ldots, \lambda_{10}, \lambda_{11}, \ldots, \lambda_{50}) = (\underbrace{5, 4.5, \ldots, 0.5}_{10}, \underbrace{0.1, 0.1, \ldots, 0.1}_{40})$$

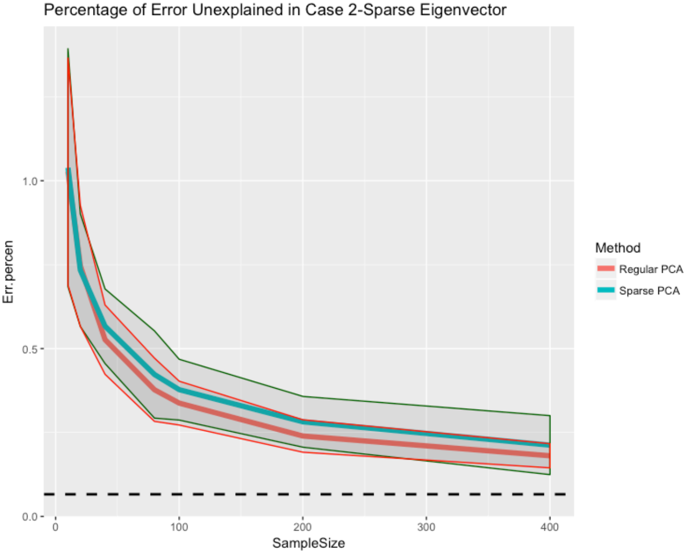
- Eigenvectors: 50 orthonormal vectors, on average, only 8.64 elements in each eigenvector are non-zero.
- Sample Size, $n = 10, 20, 40, 80, 100, 200, 400$
- Number of repetitions, $np = 20$

Case 2: Sparse Eigenvector, PVE





Case 2: Sparse Eigenvector, PEU



Thank you