A Comparison of Stochastic Methods for PCA as Applied to Streaming Facial Recognition*

Corbin Rosset
Johns Hopkins University
3400 N. Charles St.
Baltimore, MD 21218, USA
crosset2@jhu.edu

Edmund Duhaime
Johns Hopkins University
3400 N. Charles St.
Baltimore, MD 21218, USA
eduhaim1@jhu.edu

Abstract

This study applies a variety of stochastic and incremental techniques for principle component analysis (PCA) to quickly learn maximally informative linear subspaces on a streaming version of the Yale Face Data Set B. We compare the performance of a K-nearest neighbor classifier and a bagged tree learner on the best subspaces learned by each of: Stochastic Power Method (SPM), Incremental PCA (IPCA), Matrix Stochastic Descent (MSG), and Sparse PCA (SPCA). We compare and contrast the theoretical properties such as correctness, space, and iteration complexity. In all cases, the memory required to store a subspace capable of achieving accuracy similar to that of the baseline classifier was 2-4% of the size of the input data. These algorithms play an important role in improving the scalability of facial recognition in a streaming setting.

1 Introduction

The premise of subspace learning is to map a data matrix $X \in \mathcal{R}^{d \times n}$ of n examples each of dimension d to a k dimensional subspace, $k << d$ that preserves maximal "information". The motivation is that most data sets capture redundant, verbose, or noisy features that dilute the true parameters behind the data.

*We would like to thank Dr. Raman Arora for his work on stochastic algorithms for manifold learning.

2 PCA and its Stochastic Variants

introduce dual but equivalent concepts of reconstruction error and variance retention

state the empirical covariance matrix is d x d too big to be in memory

show how to solve for PCA using power method if true covariance were known

2.1 Stochastic Power Method

notes from class and relation to power method

2.2 Incremental PCA

notes from class

2.3 Matrix Stochastic Gradient

notes from class

2.4 Sparse PCA

notes from class

3 K-NN for High-Dimensional Data

discuss johnson lindenstrauss lemma popularly applied to proximity problems to preserve pairwise distances in a lower dimensional space

4 Experiments

NED: describe the contents of the data, how we removed bad images, just like proposal...

5 Results

NED: pictures of 1) reconstructed faces with 5, 10, 15... principle components 2) table of the best dimension and associated accuracy achieved by

Type of Text	Font Size	Style
paper title	15 pt	bold
author names	12 pt	bold
author affiliation	12 pt	
the word “Abstract”	12 pt	bold
section titles	12 pt	bold
document text	11 pt	
abstract text	10 pt	
captions	10 pt	
bibliography	10 pt	
footnotes	9 pt	

Table 1: Font guide.

each algorithm 3) some examples of “eigenfaces” (the top five principle components shown as images) as well as their thresholded versions. You’ll find these in the figures directory. 4) graphs of accuracy versus number of principle components for some of the algorithms (choose the better looking ones)

Captions: Provide a caption for every illustration; number each one sequentially in the form: “Figure 1. Caption of the Figure.” “Table 1. Caption of the Table.” Type the captions of the figures and tables below the body, using 10 point text.

6 Conclusion

all algorithms for pca should learn the same subspace, up to small rotations and scalings...

discuss applications of streaming subspace learning

there are extensions for incremental kernal pca and dealing with the case of missing data...

Acknowledgments

Do not number the acknowledgment section.

References

the following are examples

Alfred V. Aho and Jeffrey D. Ullman. 1972. *The Theory of Parsing, Translation and Compiling*, volume 1. Prentice-Hall, Englewood Cliffs, NJ.

American Psychological Association. 1983. *Publications Manual*. American Psychological Association, Washington, DC.

Association for Computing Machinery. 1983. *Computing Reviews*, 24(11):503–512.

Ashok K. Chandra, Dexter C. Kozen, and Larry J. Stockmeyer. 1981. Alternation. *Journal of the Association for Computing Machinery*, 28(1):114–133.

Dan Gusfield. 1997. *Algorithms on Strings, Trees and Sequences*. Cambridge University Press, Cambridge, UK.