SHIV NADAR

INSTITUTION OF EMINENCE DEEMED TO BE UNIVERSITY

DELHI NCR

EED305- RESEARCH PROJECT REPORT

Dynamic Behavior of the Whitening Process

Tianping Chen, Senior Member, IEEE, and Qin Lin

Report Submitted by (Group 22):

Aditya A (2010110037)

Karra Sanjana Reddy (2010110333)

Shlok Sudhir Kamat (2010110595)

Abstract:

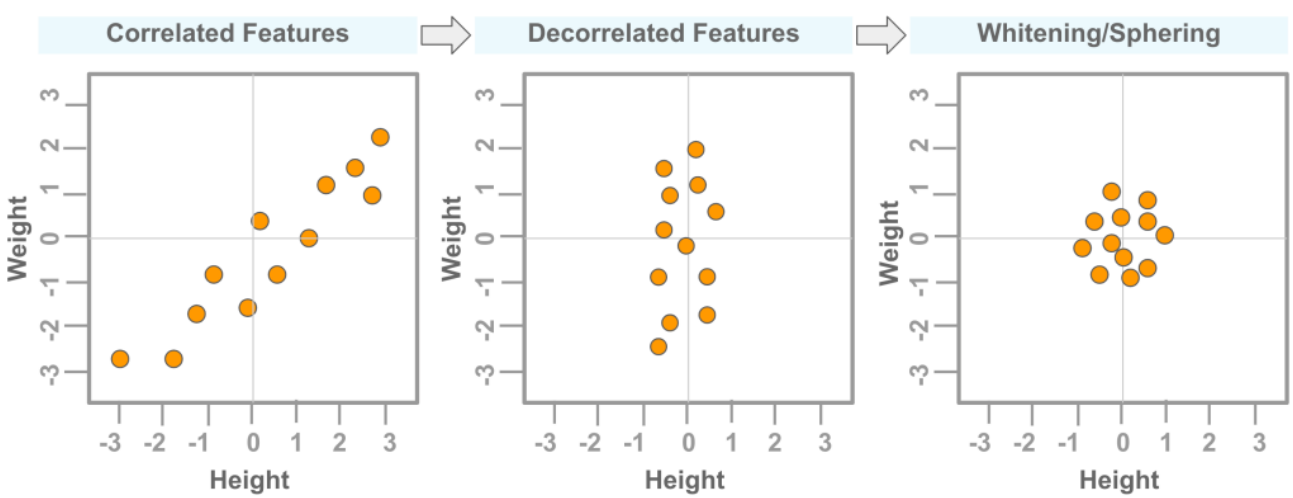
In this letter, we present an insightful view of whitening processes (or orthogonalization processes), which occur frequently in signal processing and neural network applications. For such processes, we show the dependence of the limiting solution on the initial matrix and the precise convergence rate. Simulation results corroborate our theoretical analysis.

Introduction:

In many signal processing and neural network computing applications, the requirement to spatially whiten observed signals often arises. One such application is the “blind” separation of signals. In general, an observed (recorded) time series comprises both the signal we wish to analyse and a noise component that we would like to remove. Noise or artefact removal often consists of a data reduction step (filtering) followed by a data reconstruction technique (such as interpolation). However, the success of the data reduction and reconstruction steps is highly dependent upon the nature of the noise and the signal. We explore recovering original signals by solving the “The cocktail party problem” by employing the statistical technique of Independent Component Analysis (ICA). Furthermore, we try to implement image compression using the SVD matrix decomposition.

The Whitening Transformation :

A whitening transformation or sphering transformation is a linear transformation that transforms a vector of random variables with a known covariance matrix into a set of new variables whose covariance is the identity matrix, meaning that they are uncorrelated and each have variance.The transformation is called “whitening” because it changes the input vector into a white noise vector.”Whitening or Sphering is a data pre-processing step. It can be used to remove correlation or dependencies between features in a dataset.



The mathematical definition is, suppose X is a random (column) vector with non-singular covariance matrix Σ and mean 0 . Then the transformation Y = W x X with a whitening matrix W satisfying the condition $W^T W = \Sigma^{-1}$ yields the whitened random vector Y with unit diagonal covariance.

There are infinitely many possible whitening matrices W that all satisfy the above condition. This gives rise to the various methods to calculate the whitening matrix. The various methods have their merits and demerits with varying conditions.

Commonly used choices are $W = \Sigma^{-1/2}$ (Mahalanobis or ZCA whitening), W = L T where L is the Cholesky decomposition of Σ^{-1} (Cholesky whitening), or the eigen-system of Σ (PCA whitening).

Independent Component Analysis:

Independent component analysis (ICA) is a statistical and computational technique for revealing hidden factors that underlie sets of random variables, measurements, or signals. The non-Gaussianity family of ICA algorithms, motivated by the central limit theorem, uses kurtosis and negentropy.

Independent Component Analysis

(Hérault and Jutten, 1984-1991)

- Observed data $x_i(t)$ is modelled using hidden variables $s_i(t)$:

$$x_i(t) = \sum_{j=1}^{m} a_{ij} s_j(t), \qquad i = 1 \ldots n \tag{1}$$

or as a matrix decomposition

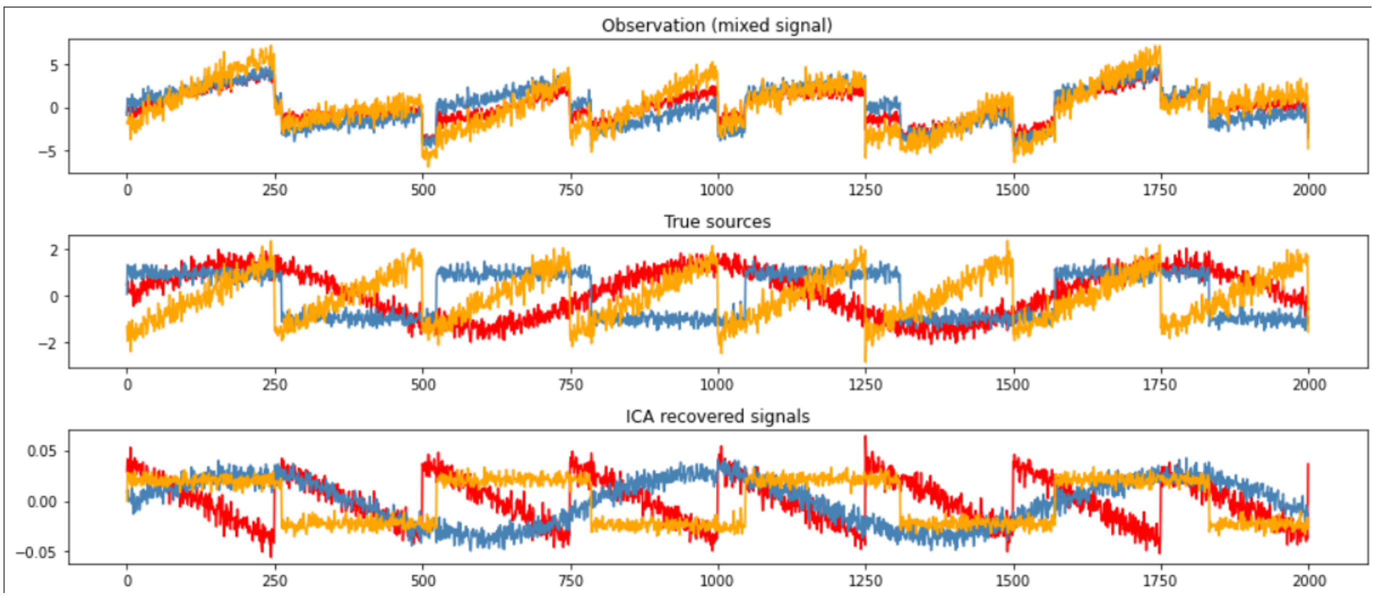
$$\mathbf{X} = \mathbf{AS} \tag{2}$$

- Matrix of a_{ij} is constant parameter called "mixing matrix"
- Hidden random factors $s_i(t)$ are called "independent components" or "source signals"
- Problem: Estimate both a_{ij} and $s_j(t)$, observing only $x_i(t)$
  - Unsupervised, exploratory approach

An excellent example of ICA usage is the classic cocktail party problem. Consider a cocktail party where many people are talking at the same time. If a microphone is present then its output is a mixture of voices. When given such a mixture, ICA identifies those individual signal components of the mixture that are unrelated. Given that the only unrelated signal components within the signal mixture are the voices of different people, this is precisely what ICA finds.

Our Implementation of BSS in Python using FastICA:

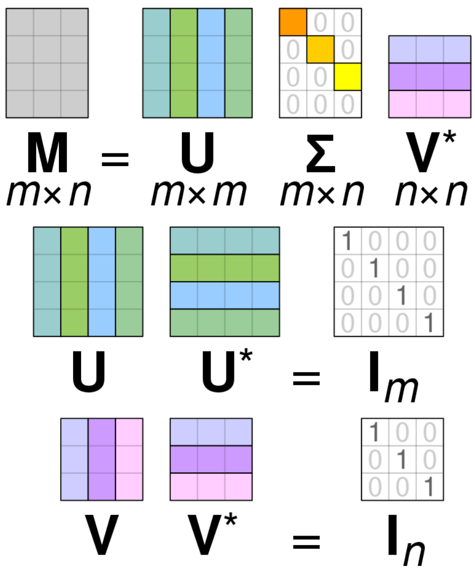
For a more palatable and practical demonstration, a python script was made to create a random mixture of three distinct signals and ICA was performed on the result. The ICA algorithm was implemented using python's builtin library. The source signals, and the result along with the mixture are displayed as a graph for comparison. Say, for a real life example, imagine three instruments playing simultaneously and three microphones recording the mixed signals. ICA is used to recover the sources, i.e. what is played by each instrument.

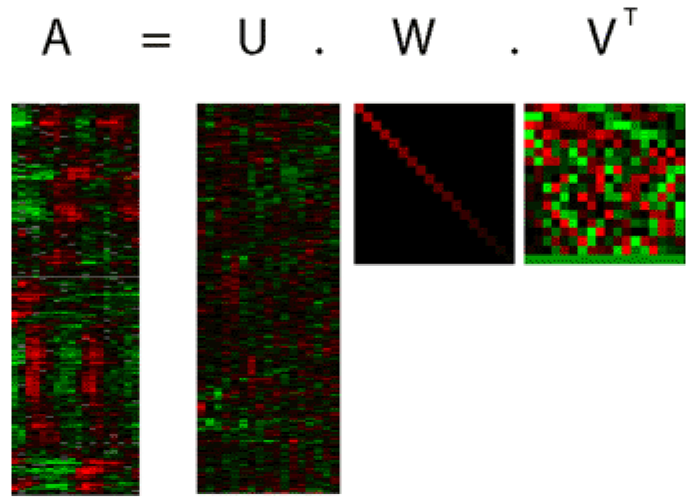


As it can be observed from the above graphs, the ICA algorithm was able to separate the different source signals from the mixed source.

The Singular Value Decomposition:

The singular value decomposition (SVD) is among the most important matrix factorizations of the computational era, providing a foundation for nearly all of the data methods in this book. We will use the SVD to obtain low-rank approximations to matrices and to perform pseudo-inverses of non-square matrices to find the solution of a system of equations. Another important use of the SVD is as the underlying algorithm of principal component analysis (PCA), where high-dimensional data is decomposed into its most statistically descriptive factors. SVD/PCA has been applied to a wide variety of problems in science and engineering.

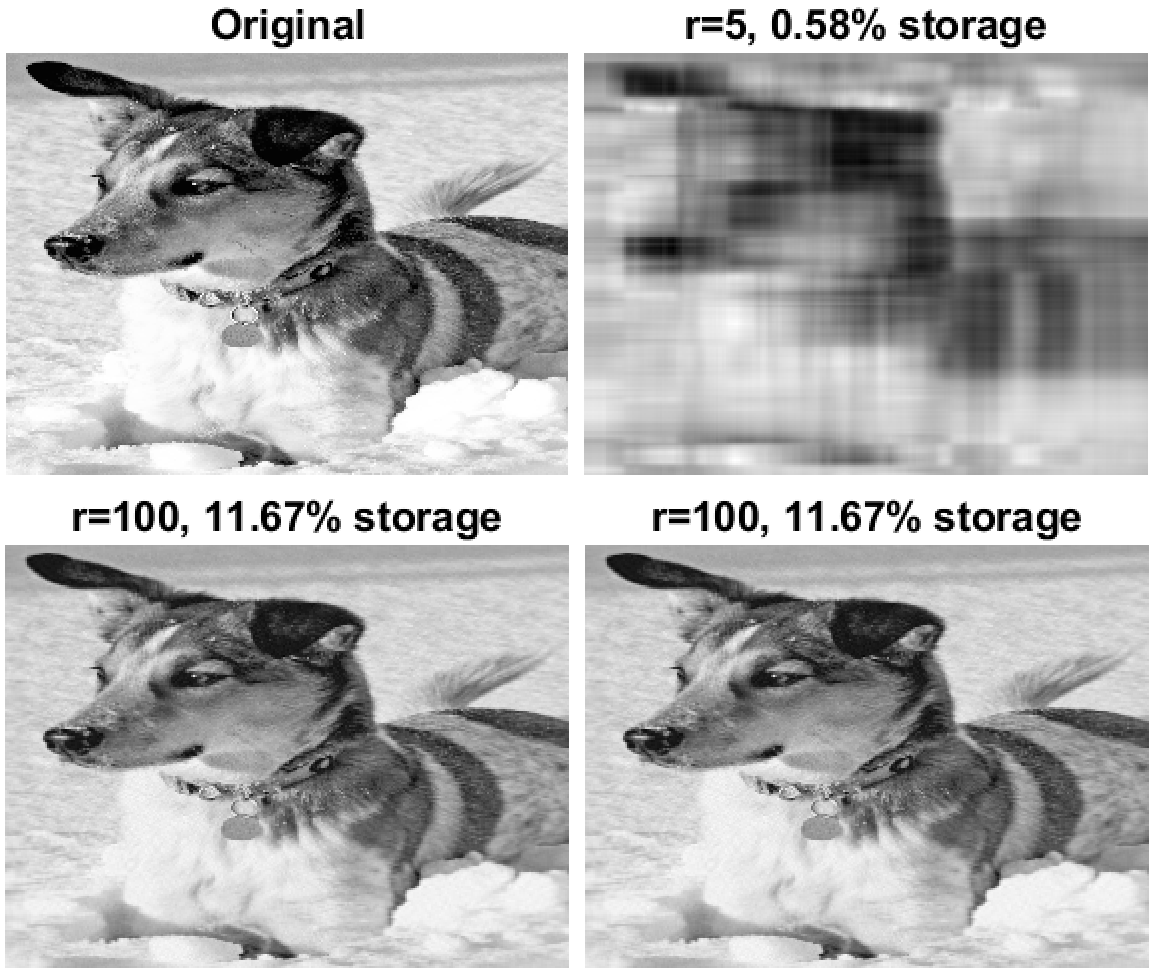




In many domains, complex systems will generate data that is naturally arranged in large matrices, or more generally in arrays. For example, a time-series of data from an experiment or a simulation may be arranged in a matrix with each column containing all of the measurements at a given time. If the data at each instant in time is multidimensional, as in a high-resolution simulation of the weather in three spatial dimensions, it is possible to reshape or flatten this data into a high-dimensional column vector, forming the columns of a large matrix. Similarly, the pixel values in a grayscale image may be stored in a matrix, or these images may be reshaped into large column vectors in a matrix to represent the frames of a movie. Remarkably, the data generated by these systems are typically low rank, meaning that there are a few dominant patterns that explain the high-dimensional data. The SVD is a numerically robust and efficient method of extracting these patterns from data.

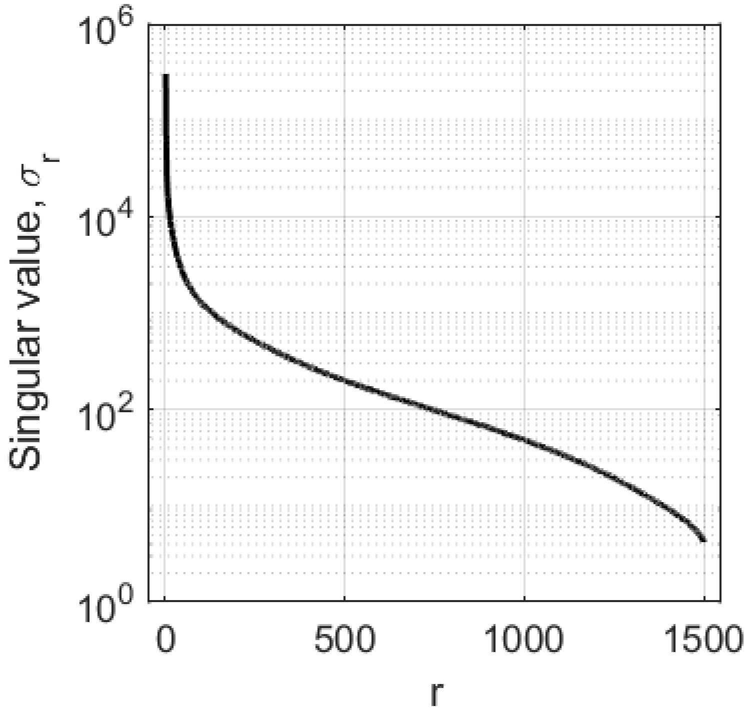
Image Decomposition using SVD:

In continuation with our methodology, a well known industry standard use case for SVD was found and implemented. Image compression is such an application which holds tremendous importance as the need for more efficient data storage and retrieval are required. SVD refactors the given digital image into three matrices. Singular values are used to refactor the Image and at the end we get Image represented with a smaller set of values,Hence reducing storage space required by Image. SVD technique is usually featured as an Algebraic feature. Algebraic features usually represent Intrinsic properties It is observed that as matrix coefficients move towards Y-axis , rank of Image matrix increases which indicates that maximum energy is concentrated with only the first few coefficients.



These are images of 'Mort', a pet dog, whose image was compressed by applying the SVD algorithm. We can clearly see that with only 100 iterations/ranks of the SVD matrix, the viewer can make out that the image is a dog. This method has direct application in digital watermarking where one can encode data in images.

The low rank nature of the images is exploited by the SVD algorithm to represent the image in a smaller size than originally required. The following graph shows the relative weights of the Singular values. This gives a more mathematical view of the reduced dataset. As can be observed, most of the image data can be represented using only a few eigen values and eigen vectors.



Conclusion:

The research paper elucidated the working and various applications of the whitening transformation including its underlying mathematical principles. Importance of whitening transformation in machine learning applications was also investigated. To further understand this, we also looked at Singular Value decomposition and its applications in image processing and dataset manipulation. A deep dive into image compression using SVD was also done and simulated in MATLAB to validate results. This also led to looking into Component Analysis techniques like Independent Component Analysis(ICA) and Principal Component Analysis(PCA). Blind source Separation techniques involving ICA were also investigated and simulated to verify the results. Even though the whitening transformation is based on very old mathematical concepts, it is still relevant today because of the above elucidated applications in emerging fields.