Principal Component Analysis

A neural network for Robust Principal Component Analysis

Eric, Mirko, Simon, Aron, Bastian

January 29, 2021

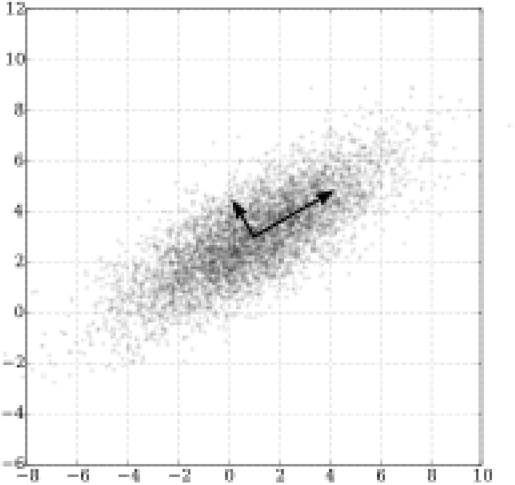
Table of Contents

Table of Contents

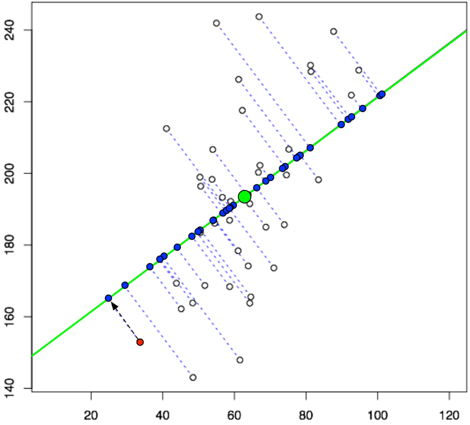
Intuition

- Observing a set of points in $\mathbb{R}^n$
- Think of fitting an n-dimensional ellipsoid to the data
- Each axis of the ellipsoid represents a principal component
- Used for dimensionality reduction
- To group the main characteristics of the data
- Makes further calculations more efficient

Axes of an Ellipsoid



Dimensionality Reduction



Mathematical Approach

- Observing a set of points in $\mathbb{R}^n$
- Sequence of n direction vectors
- i^{th} vector is direction of a line that best fits the data
- i.e. minimizes average squared distance from points to line
- And being orthogonal to first $i-1$ vectors
- Directions form orthonormal basis of $\mathbb{R}^n$
- Eigenvectors of datas Covariance Matrix

Robust PCA

- Modification of PCA for matrices with faulty entries
- Matrix might be corrupted by imprecise measurements
- Observe matrix $M \in \mathbb{R}^{m \times n}$
- Find decomposition $M = L + S$ where
- L is of low rank
- S is sparse (lots of entries are zero)
- Purpose of S is to filter out corrupted entries

Table of Contents

Principal Component Pursuit

- There are several state of the art algorithms for solving RPCA
- One of them is the so called Principal Component Pursuit (PCP)
- Solves PCA exactly under certain constraints
- Using Singular Value Decomposition and Langrange Multiplier
- Computationally demanding
- Impractical in some applications
- Due to exactness later used for Testing

Algorithm

- For $\|M\|_* := \sum_i \sigma_i(M)$
- and $\|M\|_1 = \sum_{ij} |M_{ij}|$
- Minimize $\|L\|_* + \|S\|_1$ where $L + S = M$

Table of Contents

Introducing Denise

- Feed forward deep neural network
- Trained on synthetic Dataset
- Solve RPCA for n-by-n symmetric semi-definite matrices
- Learns the decomposition map $M \to L + S$ directly
- Calculates the desired decomposition instantaneously
- Therefore saves valuable time

Denise

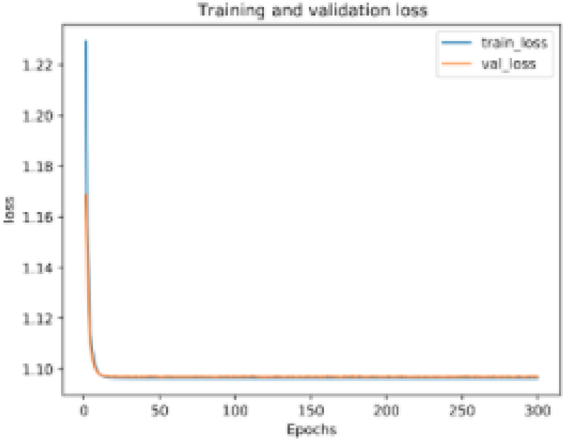
- Input: symmetric positive semi-definite matrix $M \in \mathbb{R}^{n \times n}$
- Goal: $M = L + S$, Note: L is assumed to be symmetric
- By Cholesky decomposition $M = UU^T + S$
- Lossfunction to be minimized is $\|UU^T - M\|_1 = \|S\|_1$
- Input matrix is transformed into a vector and reduced due to symmetry
- Network consists of three hidden layers and output layer
- Each hidden layer has ReLU-activation function and $n/2$ nodes

Table of Contents

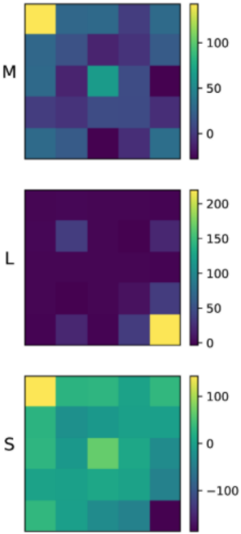
Testing

- Once trained Denise is tested on several matrices
- Portfolio Correlations from Dax 30
- Correlation matrix of personality features
- PCP is used as benchmark

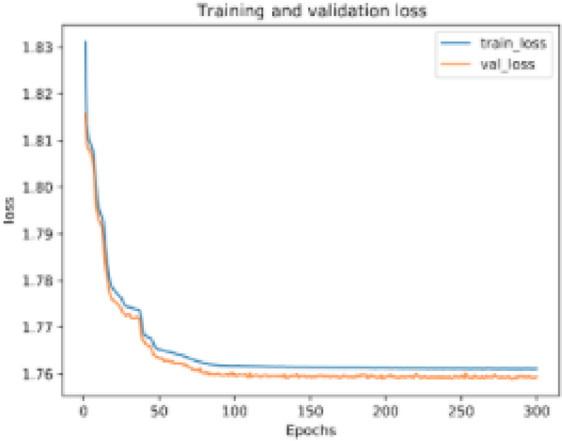
Finance Data



Finance Data



Personality Data



Personality Data