Eigenvectors and eigenvalues computation using Neural Networks

Raghvendra Mishra (ME21B1075) and Ayush Agarwal (ME21B1076)

13th April, 2023

Problem Definition

- Creating an algorithm to find the eigenvalues and their corresponding eigenvectors of a symmetric, positive definite matrix.

Symmetric Matrix: $A^T = A$

Positive definite: $z^T Az > 0$ for every non-zero column, z, of A

Problem Definition

- Creating an algorithm to find the eigenvalues and their corresponding eigenvectors of a symmetric, positive definite matrix.
- A neural network based approach will be developed to compute the eigenvectors.

Motivation for the project

Eigenvalues and eigenvector calculation is an interesting computational problem.

Motivation for the project

Eigenvalues and eigenvector calculation is an interesting computational problem.

It has uses in machine learning algorithms for Google page indexing, data analysis, clustering analysis, etc.

Motivation for the project

Eigenvalues and eigenvector calculation is an interesting computational problem.

It has uses in machine learning algorithms for Google page indexing, data analysis, clustering analysis, etc.

To provide a more **Efficient and Accurate** way of computation is something this paper strives for and hence, is trying to achieve said goal using neural network.

Mathematical Methods Used

The proposed neural network model:

$$\frac{dx(t)}{dt} = -x(t) + f(x(t))$$

for $t \geq 0$, where

$$f(x) = [x^T x A + (1 - x^T A x) I] x$$

and $x = (x_1, x_2, \ldots x_n)^T \in R^n$ represent the state of the network.
This is a class of **recurrent neural network**.

Mathematical Methods Used

The neural network is a nonlinear differential equation.

Mathematical Methods Used

The neural network is a nonlinear differential equation.

Generally, it is not easy to solve nonlinear differential equation, but the authors have used the property of A being symmetric to establish a better presentation of the solution of the neural network model
Since A is symmetric matrix, then there exists an orthonormal basis of R^n composed by eigenvectors of R^n. Let $\lambda_i (i = 1, \ldots, n)$ be eigenvalues of A and $S_i (i = 1, \ldots, n)$ are some constants.

Mathematical Methods Used

Then, solution of network starting from $x(0)$ can be represented in terms of a set of orthogonal eigenvectors.

$$x(t) = \Sigma_{i=1}^{n} \sqrt{\frac{x(0)^T x(0)}{\Sigma_{j=1}^{n} z_j^2(0) e^{2x(0)^T x(0)(\lambda_j - \lambda i)t}}} z_i(0) S_i$$

Using this representation of the network, we set a few required conditions for the solution of the neural network $x(t)$ to converge to eigenvectors corresponding to the largest eigenvalue.

Computational results from paper

These are few of the computer simulation results achieved by the paper. The simulation will show that the proposed network can calculate the eigenvectors corresponding to the largest and smallest eigenvalues of any symmetric matrix. Symmetric matrix can be randomly generated in a simple way. Let Q be any randomly generated real matrix, define:

$$A = \frac{(Q^T + Q)}{2}$$

Computational Results

First, a 5 $\times$ 5 symmetric matrix A is generated as

$$A = \begin{bmatrix} 0.7663 & 0.4283 & -0.3237 & -0.4298 & -0.1438 \\ 0.4283 & 0.2862 & 0.018 & -0.2802 & 0.1230 \\ -0.3237 & 0.0118 & -0.9093 & -0.4384 & 0.7684 \\ -0.4298 & -0.2802 & -0.4384 & -0.0386 & -0.1315 \\ -0.1438 & 0.1230 & 0.7684 & -0.1315 & -0.4480 \end{bmatrix}$$

The paper gets the estimated values of λ_{max} and λ_{min} are 1.2307 and 1.5688, accurate to the precision of 0.0001. For the minimum we feed $-A$ into the network and get 1.5688, which is accurate with just the sign flipped. The true maximum and minumum values computed by MATLAB are:

maximum eigenvalue	λ_{max} = 1.2307
minimum eigenvalue	λ_{min} = -1.5688

Computational Results

Estimated eigenvectors are:

$$\xi_{max} = \begin{bmatrix} 1.0872 \\ 0.6264 \\ -0.0809 \\ -0.4736 \\ -0.0472 \end{bmatrix} \qquad \xi_{min} = \begin{bmatrix} 0.1882 \\ 0.0600 \\ 1.3209 \\ 0.3697 \\ -0.8446 \end{bmatrix}$$

By feeding the network with -A, it gets that ξ_{min}, an estimation to the desired eigenvector, as well as the magnitude of the smallest eigenvalue, 1.5688, which is an accurate estimation just with the sign flipped. The generated eigenvector also corresponds to the target eigenvalue.

Data Used

There's still some study needed to understand how the neural network has been trained by the author's of the paper since the training data has not been mentioned in the paper.

Objective of the Project

To get a better understanding of how different neural networks work and how to simplify network models using Numerical Methods.

THANK YOU!