Lab Assignment-4

1. Write a python function for finding all eigenvalues and eigenvectors of a square matrix using QR method. Find all the eigenvalues and eigenvectors of the following matrix:

$$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 4 \end{bmatrix}$$

   Compare the results with the inbuilt python function by calculating their normed difference.

2. Write a function from scratch to find the Cholesky factorization of the following matrix:

$$A = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 4.25 & 2.75 \\ 1 & 2.75 & 3.5 \end{bmatrix}$$

   The function must raise an error if the given matrix is not symmetric positive definite. Check it by a randomly generated square matrix of size n.

3. Solve the following system of linear equations using Cholesky factorization:

   6x + 15y + 55z = 76
   15x + 55y + 225z = 295
   55x + 225y + 979z = 1259

4. **Feature Normalization and Model Performance Comparison**

   You are working on a classification task where you have a dataset with m samples and n features. Your goal is to train a machine learning model to classify the samples accurately. However, before training the model, you want to investigate the impact of feature normalization on model performance.

   Tasks:

   a. Data Preprocessing: Split the dataset into training and testing sets using a suitable ratio (e.g., 80% training, 20% testing).
   b. Feature Normalization: Implement two methods for feature normalization:

i. L1 Normalization: Scale each feature by dividing it by the L1 norm of the feature vector.
ii. L2 Normalization: Scale each feature by dividing it by the L2 norm (Euclidean norm) of the feature vector.
iii. Create two sets of datasets: one with L1-normalized features and one with L2-normalized features.

c. Model Training and Evaluation:
   i. Train a SVM classifier on both the unnormalized dataset and each of the two normalized datasets.
   ii. Evaluate the models on the testing dataset using the following metrics: Accuracy, Precision, and Recall.
d. Performance Comparison:
   i. Compare the model performance (accuracy, precision, and recall) between the unnormalized dataset and the two normalized datasets.
   ii. Analyze and discuss the impact of feature normalization on model performance. Consider aspects like time required for model training and generalization.

Note:

- Use appropriate libraries and functions for model training and evaluation. Also, use inbuilt function for normalizing the data.
- Clearly report and visualize the results for a comprehensive analysis.
- Use the following snippet of code for loading the dataset:

```
from sklearn.datasets import load_breast_cancer
# Load the Breast Cancer dataset
data = load_breast_cancer()
X = data.data
y = data.target
```

and the following snippet for loading the SVM classifier:

```
from sklearn.svm import SVC
# Create an SVM classifier
clf = SVC(kernel='linear')
# Train the classifier
clf.fit(X_train, y_train)
# Make predictions on the test set
y_pred = clf.predict(X_test)
```